Understanding and exploiting livestock grazing behaviour for ecosystem service delivery

Benjamin Philip Roberts

A thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy (PhD)

2020

Institute of Biological Environmental and Rural Sciences

Aberystwyth University

DECLARATION OF WORK

This work has not previously been accepted in substance for any degree and is not being concurrently submitted in candidature for any degree.

Candidate name:

_____Benjamin Philip Roberts_____

Signature: 

Date: 23 /08 /2020

STATEMENT 1

This thesis is the result of my own investigations, except where otherwise stated. Other sources are acknowledged explicitly in the text. A bibliography is included.

Signature: 

Date: 23 /08 /2020

STATEMENT 2

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signature: 

Date: 23 /08 /2020

Word Count of Thesis (excluding appendices and references): 52,125

ACKNOWLEDGEMENTS

The level of technological development undertaken in this project would not have been possible without the considerable input and advice offered by my many secondary supervisors in the Department of Computer Science across the lifetime of the study. I would like to thank Dr Mark Neal (now at Ystumtec Ltd, but formerly of Aberystwyth University) for ultimately undertaking the hardware technical design and construction of the UAVRTS and GNSS collars. I am particularly grateful for his patience as I consistently took out his new shiny equipment, tested it, inevitably broke it, then returned it for modification. The same is also true of Dr Neal Snooke who spent many hours in the robotics shed helping me try (and eventually succeed) to construct the fixed wing UAVs. I would also like to thank him for realising my ideas of the Beferrer/Schedsplitter application by writing the JAVA code for them based on my ever-expanding specifications. Finally, I am also indebted to Dr Fred Labrosse for his support in the latter stages of my PhD, particularly through writing the UAV speed estimation code that was a crucial revision required for the UAVRTS section of the study to be published. Additional thanks in the computer science department must also go to Dr Pete Todd (for his soldering skills in tricky situations), Mr Barry Thomas (for cutting the plywood reinforcements for the UAV), and Dr Matt Gunn (for procuring the Micasense Rededge cameras). I would like to thank the Earth Observation and Ecosystem Dynamics Research Group within the university (particularly Dr Pete Bunting and Dr Andy Hardy) for their advice regarding the remote sensing aspects of the project, and the use of their equipment and software which was vital in the initial part of the project. I would also like to thank all the technical team (past and present) (Iolo, Hannah, Eirian, Gareth, John and Tom) at the Pwllpeiran Upland Research Centre whose sheep gathering/UAV launching skills/remote desktop restarting services were required on several occasions throughout the project. Thanks must also go to Mike Yearby (Comp Sci undergraduate student), James Hinkley (MSc student in remote sensing and GIS), and Alex Noble (IBERS MSc student) whose willingness to conduct their dissertations on aspects surrounding my project, allowed me better focus on which pathways to take. Final thanks in the university must go to my primary supervisor Dr Mariecia Fraser whose guidance and advice has been invaluable over the past four years, and who deserves a ‘long-service’ award for sticking as a supervisor throughout the whole PhD.

Thanks, must be given to my friend Tristan Curteis, whose expertise as a statistician was utilised to undertake the statistical analysis for the UAVRT section. Dr Ian Johnstone (at

RSPB) and Dr Ieuan Joyce (Elan valley trust) are also thanked for their key inputs in project meetings, as well as the technical teams at RSPB Lake Vyrnwy and Elan valley whose assistance with the livestock on the study sites was crucial to project success.

Finally, thanks must be given to Pippa my girlfriend and Rosie the Jackhuahua for academic support and sharing of PhD stresses when the going was tough, as well as my parents for fieldwork and proof-reading support. Other friends and family are also thanked for their support, and who will no doubt be pleased that the social obligation to ask on PhD progress will no longer be obligatory.

ABSTRACT

The use of grazers for vegetation control and restoration is common practice in agri-environment schemes and in wider habitat management exercises. However, there is a fine balance in achieving the optimum grazing intensity on a given site, with many habitats currently under-or over-grazed. In order to achieve best practice, it is important to understand the animal-environment interactions occurring on a given site. This refers both to the effect the animals are having on their surrounding environment, and also how the animals are being influenced by their environment. This latter point is important because knowledge of which environmental variables are driving grazer distribution can facilitate manipulation to allow more targeted grazing regimes. This study aimed to develop a standardised workflow that begins with initial site profiling and ends with the capacity to provide informed recommendations on future management interventions at large-scale study sites. The ability of recent and emerging technologies to address the issues of large site data gathering were the foundation for the workflow success. A combination of custom developed unmanned aerial vehicles (UAVs) and Global Navigation Satellite System (GNSS) tracking collars were used to collect data relating to animal movement and environmental factors at two sites (72ha and 112ha) in Mid-Wales. A series of analytical approaches were developed and employed, including simple practitioner methods which could be used by land managers, as well as more complex statistical techniques which can address academic hypotheses relating to the site. These included a bespoke livestock unit (LSU) estimator based on GNSS recordings, individual and combined home range analyses, GIS visual interpretation, customised software for estimating livestock behaviour, and species distribution modelling. Results from practitioner methods revealed grazing pressure to be uneven at both sites, with the majority (55.6% and 56.7% respectively) of areas at both sites being under-grazed. An ensembled modelling approach revealed different environmental variables to be affecting sheep distribution at each site, thereby validating the need for site specific approaches to recommending prescriptions. The amount of data produced, and the ability to utilise different combinations of newly developed/existing tools (e.g. behavioural inference, species distribution modelling, home range estimation) to understand how distribution and grazing behaviour is being driven on a specific site, lends itself as an effective tool for being able to recommended targeted management interventions.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

AGL	Above Ground Level
AIC	Akaike Information Criterion
AKDE	Autocorrelated Kernel Density Estimation
AMSL	Above Mean Sea Level
ANN	Artificial Neural Network
ASTER	Advanced Spaceborne Thermal Emission and Reflection Radiometer
AUC	Area Under the Curve of the receiver-operating characteristic
BRT	Boosted Regression Trees
BVLOS	Beyond Visual Line-Of-Sight
CAA	Civil Aviation Authority
CAP	Common Agricultural Policy
CEP	Circular Error Probable
CIR	Colour-Infrared
CRP	Calibrated Reflectance Panel
CRS	Coordinate Reference System
DEM	Digital Elevation Model
DLS	Downwelling Light Sensor
DSM	Digital Surface Model
DTM	Digital Terrain Model
ENM	Ecological Niche Models
EPP	Expanded Polypropylene
ESA	European Space Agency
FBWA	Fly-By-Wire-A
FMD	Foot and Mouth Disease
GAM	General Additive Model
GDAL	Geospatial Data Abstraction Library
GHG	Greenhouse Gas
GIS	Geographic Information Systems
GME	Geospatial Modelling Environment
GNSS	Global Navigation Satellite System
GSD	Ground Sampling Distance
GUI	Graphical User Interface
IBERS	Institute of Biological, Environmental & Rural Sciences
IMU	Inertial Measurement Unit
KDE	Kernel Density Estimators
LFA	Less Favoured Areas
LSU	Livestock Unit
MA	Millennium Ecosystem Assessments
MARS	Multivariate adaptive regression splines
MAXENT	Maximum Entropy
MCP	Minimum Convex Polygon
MLP	Multi-Layer Perceptron
MOD	Ministry of Defence
MTBF	Mean Time Between Failure

NDVI	Normalised Difference Vegetation Index
NDWI	Normalised Difference Water Index
NIR	Near-Infrared
NRW	Natural Resources Wales
PID	Proportional–Integral–Derivative
PPK	Post-Processed Kinematic
RBF	Radial-Basis Function
RC	Radio-Control
RE	Red-Edge
RF	Radio-Frequency
RFID	Radio-Frequency Identification
RSPB	Royal Society for the Protection of Birds
RTH	Return-To-Home
SAC	Special Areas of Conservation
SDM	Species Distribution Modelling
SEM	Standard Error of the Mean
SfM	Structure from Motion
SFP	Single Farm Payment
SPA	Special Protection Area
SRTM	Shuttle Radar Topography Mission
SSSI	Sites of Special Scientific Interest
SVM	Support Vector Machine
TSS	True Skill Statistic
TTFF	Time To First Fix
UAA	Utilizable Agricultural Area
UAV	Unmanned Aerial Vehicle
UBEC	Universal Battery Eliminator Circuit
UD	Utilisation Distributions
VIF	Variance Inflation Factor

CHAPTER 1

INTRODUCTION

1.1 The global impact of grazing

Globally, grazing land is the largest and most diverse land type in the world, with more than half the Earth's land surface being grazed by domestic and wild ungulates (Follett & Reed 2010). From a domestic perspective, managed grazing is estimated to have increased 600% in geographic extent in the last 300 years, with 33 million km^2 (25% of the global land surface) now utilised for the production of meat, milk and other animal products (Asner *et al.* 2004). With regards to the benefits offered, managed grazing contributes 40% of global agricultural gross domestic product, creating income and nourishment for more than 1.3 billion people (of which, at least 800 million are food-insecure people) (Herrero *et al.* 2013). The environmental implications of grazing however in many cases have been detrimental. Soil compaction and erosion via trampling (Evans 2005), groundwater contamination (through poor manure management) (Menzi *et al.* 2010), greenhouse gas emissions (e.g. methane and nitrous oxide) (Forabosco, Chitchyan & Mantovani 2017), and a reduction of species diversity through intensification (Luoto *et al.* 2003) have all been key issues. This latter point is more pronounced when animals have greater opportunities for selective grazing (i.e. when vegetation is more botanically and structurally heterogeneous) and reflects the difference between cultivated and natural or semi-natural vegetation. Poor grazing practices can result in overbalanced selection of particular species (thereby increasing the likelihood of rarity or loss within an ecosystem), whereas good grazing practices can prevent species imbalance from occurring, and even promote floristic diversity (and therefore more wider biodiversity) (Hartnett, Hickman & Walter 1996). Understanding the dynamic between animals and the vegetation they are grazing in natural or semi-natural vegetation, is therefore central to management for productive and/or environmental gain. This is particularly true in marginal areas (i.e. where land has little or no agricultural or industrial value), where meeting multiple goals is frequently required when managing vegetation (Kang *et al.* 2013). Within the United Kingdom, marginal land is predominately encompassed with upland areas (Acs *et al.* 2010).

1.2 Uplands of the UK

1.2.1 The definition, and classification of UK uplands

Although there is no statutory definition of UK uplands, the Department for Environment, Food and Rural Affairs (2011) 'Upland policy review' recognises the Government's adoption of the definition of land categorised as "Less Favoured Areas (LFA)". Defined in accordance

with EU Directive 75/2767 – LFAs were the European designation for areas with natural and socio-economic disadvantages, which in the UK largely corresponds to areas of upland farming systems. LFAs therefore comprise the majority of the uplands of England, Scotland, Wales, and Northern Ireland. They include 53% of utilizable agricultural area (UAA) totalling 9.12 million ha within the UK, representing 17% of the UAA in England, 86% in Scotland, 80% in Wales, and 70% in Northern Ireland (Jones, Silcock, & Uetake, 2015). There are other definitions for the extent of upland areas (Bunce 1987; Averis 2004; RSPB 2007), which Bunce, Wood & Smart (2018) note have broad consensus in the general area for uplands, but with detail and comprehensive coverage lacking. Finally, though there is variation in estimation, the number of habitat classes for uplands habitats is thought to be ~45 (Bunce *et al.* 2007).

Figure 1.1 Distribution of upland land classes in Britain, derived from the ITE land classification (Bunce et al. 2007).

1.2.2 The history of human management in UK uplands

Throughout the second half of the 20th century to the start of the 21st, human management within the uplands altered considerably, with profound results on upland ecosystems. Since the start of the 1950s, improved hill pastures, species-poor acid grasslands and conifer plantations, through afforestation, improvements for agriculture and game rearing have replaced many of the natural habitats found in the uplands (Reed *et al.* 2009; Buchanan *et al.*

2017). Surviving areas of upland habitats have also seen widespread reductions in their quality and range due a number of direct and indirect human factors such as: over-grazing/under-grazing; ploughing; soil improvements (e.g. fertiliser addition); peat extraction; inappropriate burning; quarrying; and atmospheric pollution (Evans 1997; Orr *et al.* 2008). Throughout this time, agricultural policies have had a huge influence on the land management techniques imposed. From a focus of agricultural enhancement (through the 1947 Agricultural Act and the EU Common Agricultural Policy (CAP)) to meet the nation's need for food at reasonable prices for farmers and consumers (Reed *et al.* 2009), towards a decoupled direct payment attached to cross compliance with environmental and health standards, upland management has altered much during this time. The importance, and future potential of uplands for not only agricultural and commercial forestry but also ecosystem delivery has been realised and more recent government policies reflect this. DEFRA's 2011 upland policy review outlines several utility options that uplands provide: (1) improving water quality through reductions in diffuse pollution upstream; (2) managing upstream land for flood mitigation; (3) carbon storage (for example, on peat uplands); (4) biodiversity conservation; and (5) cultural ecosystem services, which include recreational benefits (DEFRA 2011).

1.3 The importance of upland peatlands

1.3.1 The characteristics of upland peatlands

Blanket bogs and raised bog peatlands represent the most widespread habitat in UK uplands, with 25% of upland areas classed as such (Bunce, Wood & Smart 2018). Peatlands are areas of land with a naturally occurring layer of peat; peat being characterised as an accumulation of partially decaying vegetative material (Hao 2019). The formation of peat is generally controlled by the balance of temperature and water as combined conditions (Moore 1987). A general provision for peat forming is that accumulating organic matter input is greater than the decomposing quantity, with the accumulating rate being determined by different combinations of these conditions, influencing the accumulation and decomposition of peat, and consequently the ecological characteristics of the peatland. The spatial and temporal changes of temperature and water primarily depend on climatic condition, with geomorphic, geological, and hydrological factors also influencing peat formation (Xintu 2008). Within a global context, the UK is within the top ten nations in the world for total peatland area, with 13% of the global blanket bog area present within the UK (Bain *et al.* 2011). As such, the

international prominence of UK peatlands makes their conservation and management of the upmost importance.

1.3.2 The ecosystem services provided by peatlands

Although there is discussion over the exact definition (Fisher, Turner & Morling 2009), 'ecosystem services' can simply be considered as the 'the benefits that people obtain from ecosystems' (Kimmel & Mander 2010). The concept is increasingly employed by policymakers, and is ever present within academic and practice debates on how to more accurately consider the value of environmental resources in decision making (Apitz 2013; Lennon & Scott 2014).

The Millennium Ecosystem Assessments (MA) provides four broad categorisations into which peatlands' ecosystem services are classified; provisioning, regulating, cultural and supporting services (MA 2005). Examples of ecosystem services that upland peatlands provide include: (1) provisioning – food, fibre, fresh water (they provide over 70% of Britain's freshwater) (Natural England 2009), and sources of energy (Tuohy *et al.* 2009); (2) cultural services – includes recreation opportunities, enjoyment of landscape aesthetics, biodiversity and cultural heritage as well as spiritual enrichment and educational experiences (Grand-Clement *et al.* 2013); (3) supporting services – defined as "*services that are necessary for the production of all other ecosystem services*" (Atkins *et al.* 2011) and include nutrient cycling, microclimate regulation, soil formation and photosynthesis (Kimmel & Mander 2010); (4) regulating services that provide benefits from regulation of natural processes. In the case of peatlands, these latter services include air quality regulation, natural hazard regulation and water regulation (including purification, flood regulation, and as a consequence regulation of natural hazards such as floods and wildfires) (Bonn *et al.* 2010; Joosten *et al.* 2014). However, perhaps the most analysed and valued of the regulating services provided by peatland is climate regulation (Joosten *et al.* 2016), with carbon sequestration considered a justification for 62% of active English peatland restoration projects in Lunt *et al* (2010).

Globally, though peatlands only cover 2-3% of the Earth surface (Evans *et al.* 2014), they are estimated to represent up to one third of the world's terrestrial carbon store (Worrall *et al.* 2010), making them the largest terrestrial store of organic carbon (Joosten *et al.* 2016). In the UK alone, peatlands store over 3200 million tonnes of soil carbon (Bain *et al.* 2011), far exceeding the total carbon stored in UK woodland vegetation of 92 million tonnes (Bonn et

al., 2010). The high volume of carbon stored means any losses in peatland area has devastating consequences. Losing just 5% of the UK peatland area would equate to the total annual UK anthropogenic greenhouse gas emissions for 1 year (Bain et al., 2011). Therefore, the extent and effect of UK upland blanket bogs on the UK's national targets to reduce its greenhouse gas (GHG) emission is profound. Blanket bogs and raised bogs are also priorities for conservation under the EC Habitats Directive. From a financial standpoint, a recent ecosystem services valuation estimated that blanket bogs are worth £226 million per year to the UK economy in terms of climate regulation and other services alone (Christie et al. 2011). This is despite their generally degraded state.

1.3.3 The state of upland peatlands

It is estimated over half of Europe's peatlands have been lost, with the largest of the these losses occurring over the last 75 years (Andersen *et al.* 2017). Currently, 1.8 million of the 2.2 million ha of peatland area in the UK is thought to be damaged in some way (Bain *et al.* 2011). The best available evidence suggests that less than 20% of UK blanket bog is in a natural or near-natural condition (Littlewood *et al.* 2014). Within the most important, nationally, and internationally protected sites (SSSIs/SACs /SPAs) only around half (58%) of the blanket bog habitat is believed to be in favourable condition. Of the remainder, only 15% is considered to be recovering as a result of restoration work (JNCC, 2009). Recognising the current and potential future value of this habitat, the IUCN UK Commission of Enquiry on Peatlands (Bain *et al.* 2011), suggested that "*A positive interim target would be to work towards having 1 million ha of peatlands in good condition or under restoration management by 2020*"-a timescale consistent with UK and international biodiversity objectives as well as commitments to tackle global climate change. This has since been updated by the IUCN UK peatland programme (2018) to include additional milestones, notably that 50% of peatlands are in good condition by 2030, and 95% by 2040. Given the scale of the restoration required, and the limited time in which to reach the milestones set, it is crucial that restoration techniques are effective and efficient.

1.4 Methods of restoring upland peatlands, and the move towards integrated management techniques

Restoration of damaged peatlands is often targeted as a means of restarting their carbon (C)-sink function. In general, restoration involves re-establishing the peatland hydrology,

biogeochemical cycling, and energy capture that promote autogenic plant succession (Gorham & Rochefort 2003). Since the start of the 1980s, drain or 'grip' blocking has been utilised as a peatland restoration method. With drain blocking, it is theorised that as the water tables rise, the thickness of the aerobic layer is reduced, resulting in an overall decrease in the rates of peat decomposition, with subsequent reduced rates of CO_2 emissions to the atmosphere. Depending on how the local plant productivity is affected and the rates of fluvial C losses, peat accumulation (a C sink) may restart, representing a net uptake of CO_2 (Green *et al.* 2011). Though drain blocking has been well established as a restoration method, there remains a great deal of inconsistency in the outcomes at different sites (Wilson *et al.* 2010). Studies of the impact peatland restoration in the uplands have so far tended to focus on water table depth and carbon efflux (Cooper *et al.* 2014; Peacock *et al.* 2015). Wilson *et al* (2010) suggest that the complex nature of peat formation means restoration through grip blocking, with the intention of rewetting alone is unlikely to be successful long-term. Though several studies carried out in Scandinavia have found that ditch blocking increased the cover of specialist bog plants such as *Eriphorum angustifolium* and *E. vaginatum* (Komulainen *et al.* 1999), Green *et al* (2017) found no evidence that drain blocking had a significant effect on the vegetation at the Migneint blanket bog in N Wales over a four year period. Furthermore, at four upland blanket peat catchments in northern Britain, Holden, Gascoign & Bosanko (2007) found that only 35% of all drain cross-sections surveyed were revegetating and infilling. Local site characteristics such as water table topography (Williamson *et al.* 2017), rainfall, degree of ditch erosion, (Armstrong *et al.* 2009), as well as peat structure and catchment size (Wilson *et al.* 2010), all influence the success of re-wetting activities.

Given that a fundamental objective of rewetting is to alter the vegetation within a habitat to something more desirable for ecosystem functioning (e.g. *Sphagnum* moss for peat formation in peatlands), manipulating the vegetation through other means is frequently explored. Typically, cutting, burning, grazing, revegetating (e.g. sphagnum seeding) (or a combination thereof) are utilised alongside rewetting procedures to promote and expedite peatland recovery (Bakker 1989; Ward *et al.* 2007; Rosenburgh 2015; Andersen *et al.* 2017).

Utilizing multiple methods in a habitat restoration project is a reasoned approach. Afterall, it is sensible to propose that the inadequacies of a single approach on a particular site may be overcome with multiple methods employed (Milligan *et al.* 2004; Parry, Holden & Chapman 2014). When considering habitat restoration more widely, the inconsistencies in results from using a singular vegetation manipulation approach (e.g. burning, cutting) in a multitude of

different UK habitats (e.g. heather moorlands, coastal sand dune systems, saltmarshes) is prevalently seen. Of all these approaches, perhaps the one most common example of detrimental mismanagement in the UK is grazing, particularly in upland areas.

1.5 The issues within habitat restoration

1.5.1 A brief history on livestock grazing in the UK, and the implications on habitat management.

It has been recognised for many years that grazing animals play a major role in modifying and shaping the environment they are in, often having a more significant influence on the productivity and structure of their vegetational environment than is generally recognised. Both historically and currently, a considerable proportion of upland vegetation is subject to grazing, particularly by sheep (*Ovis aries*) (Table 1.1). In the second half of the 20th century, the use of headage payments as direct production subsidies through the CAP led to excessively high stocking rates in many upland hill areas (Hanley, Whitby & Simpson 1999). However, generally livestock numbers within UK uplands have declined since then. For example, the total breeding flock of sheep in the UK went from a peak of ~31 million in 1998 to low of 22 million in 2008, and as of 2017 was 23.3 million (-24.8% reduction since 1998) (Bunce, Wood & Smart 2018). This has been attributed to a number of factors, including: (1) poor profitability of livestock farming; (2) the switch from headage to decoupled payments in 2003 (and subsequent formation of the Single Farm Payment (SFP)); (3) the introduction of agri-environment schemes; (4) outbreaks of livestock diseases such as Foot and Mouth Disease (FMD); and (5) more general socioeconomic factors such as an aging farmer population and growth in off-farm income leading to less demanding systems, both in labour and management (Tranter *et al.* 2007; Silcock, Brunyee & Pring 2012).

The positive consequence of this reduction is that concern of over-grazing has lessened considerably in recent years. This is exemplified by Martin *et al* (2013), who note that the upland SSSIs areas where over-grazing was a key threat had reduced from 230,000 ha in 2003 to less than 2,500 in 2013. However, it is now postulated that in many instances, under-grazing may be an increasing concern. Though the exact measured effect and extent of under-grazing in the UK is not yet defined, it is thought the spread of dominant competitor species, such as *Pteridium aquilinum* (bracken) and *Molinia caerulea* (purple moor grass; hereon referred as *Molinia*) is partially as a result of the shift in grazing intensities (Milligan *et al.* 2004; Martin 2013).

The encroachment of *Molinia* in peatlands, fens and moorlands with the UK has been viewed with considerable concern (National Trust, 2015). The ability of *Molinia* to re-translocate 75-85% of Nitrogen and Phosphorous from senescent leaves into their root and basal internodes for use in the following season, makes them highly competitive in unmanaged swards with low P and K availability and high N deposition (Hejcman *et al*. 2010). Indeed, within England, it is estimated that ~17,000-24,300 ha has ≥75% cover of *Molinia* which represents 4-6% of the total area of moorland (National Trust, 2015). Efforts to reduce *Molinia* dominance have mainly centred on utilising either cutting or grazing, or even a combination of both (Hejcman *et al*. 2010). Martin *et al* (2013) calculated that a grazing utilisation of 33% of annual leaf growth can reduce Molinia biomass/ cover, and promote increased floristic diversity. As such, practitioners and policy makers continually seek to derive the optimum balance of grazing intensity necessary to restore and maintain a habitat to a desired conservation state. One of most commonly used metrics to define optimal intensity with the UK are Livestock Units (LSU).

Table 1.1 Percentage and area of grazing animals recorded in land parcels within the sample 1 km squares (Bunce, Wood & Smart 2018).

Upland zone	**Animal**	**% upland area grazed**	**Area ('000ha)**
Uplands (England)	Cattle	4.2	65.6
	Deer	0.1	1.1
	Grouse	7.1	112.1
	Sheep	39.9	628.4
	Any grazing animal	45.3	713.1
Intermediate uplands and islands (Scotland)	Cattle	3.3	97.7
	Deer	11.8	352.5
	Grouse	2.3	69.5
	Sheep	42.5	1270.5
	Any grazing animal	45.3	1516.7

True uplands (Scotland)	Cattle	0.7	21.4
	Deer	37.4	1197.0
	Grouse	17.5	559.1
	Sheep	37.8	1209.9
	Any grazing animal	56.5	1808.5
Uplands (Wales)	Cattle	4.1	42.4
	Sheep	47.4	487.4
	Any grazing animal	48.7	500.0

1.5.2 Grazing prescriptions for habitat management-the concerns of using livestock units (LSU).

LSUs account for differences between livestock types in broad liveweight categories, as well as un-weaned young in ascribing livestock units (i.e. the figures for a ewe include lambs at foot). They are widely used within UK Agri-environment schemes (e.g. Glastir in Wales, Countryside Stewardship in England) (Natural England 2010; Harvey *et al.* 2019), and as such are a favoured system for standardising grazing prescriptions within conservation. They are not however without their concerns. Martin *et al* (2013) note that whilst they allow some comparison between different experiments, they generally hide the seasonality of grazing (i.e. animal-plant interactions will vary over the year), and they do not account for differences in breed/species in grazing behaviour. This includes the suggestion that animals may have different home ranges/hefts. The implication of this is that because LSUs assume uniformity in grazing across a given area, allowance is not being made for uneven grazing pressure. The result of this, is that on a given site, areas may be over-grazed (where conditions are preferable), and under-grazed (on lesser preferable areas). As such, it is likely that the progress of many restoration projects may be stalled, or even detrimentally affected where animals actively avoid less palatable dominant species (e.g. *Molinia*), and selectively graze the species that are intended to be conserved. This is further exacerbated when it is considered that the LSU system does not account for different site-specific characteristics (e.g. topology, climate, vegetation). As Garcia *et al* (2013) conclude, *"grazing effects depend*

on many factors, there is not a single grazing pressure that will be appropriate (1) across all sites, (2) on the same site in different years, or (3) for all fauna".

1.5.3 The importance of considering site specific characteristics, and being able to collect baseline data over a large spatial scale.

Parry, Holden & Chapman (2014) state that when it comes to blanket bog restoration projects, individual site characteristics are not appreciated sufficiently, and that practitioners should account for this variability when designing restoration strategies and monitoring impact. Additionally, when reviewing the progress and challenges in peatland restoration projects in Western Europe between 1993 and 2015, Andersen *et al* (2017) concluded that monitoring was limited to the sites under restoration (i.e. no reference sites or baselining first). Without baseline information, they caution that the data is often not sufficiently comprehensive to conduct statistical analysis. They conclude that '*monitoring ecosystem functions against baselines and references is necessary to assess "success" of restoration but is currently mostly lacking in Western Europe'*. A particular instance of where baselining is key is when existing grazing (e.g. by livestock) is already happening on a site, or where previous interventions (e.g. cutting/burning) have already taken place. Additionally, Andersen *et al* (2017) note that ground based-measurements are unrealistic for large-scale restoration projects. Given restoration of upland areas are often at a landscape scale (i.e. a whole moorland plateau or mountainside), the use of ground-based methods does not allow for micro-variations in site characteristics to be accounted for. Instead, they advocate the development of cost-effective methods (e.g. remote sensing approaches) to better quantify and monitor the target sites.

1.5.4 The opportunities using new technologies

Currently, many of the techniques used in on-ground surveying (e.g. point transect/frame methods for plants) have not changed in their approach for many decades. Whilst this does not diminish their effectiveness, their use in large-scale site surveys are typically impractical to be representative. The rapid development of remote sensing and geographic information systems in recent years has prompted ever increasing interest as to the possible ecological applications (Kerr & Ostrovsky 2003; Cohen & Goward 2004). The ability to conduct comprehensive analyses directly over large areas without the need for representative sampling, combined with the rapidly increasing resolution (<30m) at which data can be attained, offers improved possibilities for surveying environmental variables in land

management/habitat restoration exercises. The quantity of data produced is not restricted to remote sensing. Widespread use of global navigation satellite systems (GNSS) has greatly enhanced the spatial and temporal resolutions at which animals can be tracked, whilst the recent development of unmanned aerial vehicles (UAVs) has facilitated their use as tailored high-precision sensor platforms (Bhardwaj *et al.* 2016). The technological opportunities therefore to survey and track animals on large sites are numerous. However, the cost and expertise required to utilise any approaches must be considered carefully given the restrictions practitioners often have. Solutions must therefore seek to be low-cost and easy to use where possible but also capable of generating data which can be statistically rigorous when analysed.

1.6 This study

1.6.1 The motivation behind the study

Though improving the restoration of UK upland peatlands has been a key justification for this study, regardless of the habitat under restoration, there is a general need for tailored site-specific data gathering approaches, which facilitate comprehensive baseline information collection across large-scale (+50 ha) sites. The presence of such data would allow for detailed insight as to the composition of a particular site and allow the effects of any existing management exercises to be quantified and understood. Given the prevalence of livestock within upland habitats and the detrimental implications mismanagement can bring to restoration projects, emphasis must be placed on understanding the dynamics of grazing behaviour on a given large-scale site. By doing this, insight into the use of existing grazing prescription systems (such as LSUs) across different sites, can be evaluated and discussed. Finally, whilst any developed techniques must be suitable for large-scale data collection, they must also be sufficiently affordable to reflect the typical finances available for habitat restoration project. The data generated should also be comprehensive enough to allow for rigorous statistical analysis to occur, whilst consideration is also given to producing simple analytical techniques which practitioners can utilise. Through addressing these points, it will then be possible to suggest management interventions for facilitating restoration, which are tailored to a specific site. These suggestions will be evidence-based, and the result of any future interventions could then be accurately quantified against the baseline using the same data collection methods.

1.6.2 The aims and objectives of the study

This study aimed to develop a standardised workflow that begins with initial site profiling and ends with the capacity to provide informed recommendations on future management interventions at the study sites. The ability of recent and emerging technologies to address the issues of large site data gathering was the foundation for the workflow succeeding. The key research question posed was.

'Can recent developments in sensing technologies and/or processing capability be harnessed to improve the evidence-base for conservation decision-making?'

Underpinning this workflow was an investigation into existing methodologies which can be used to collect, process, and analyse the data required, as well as the development of new hardware/software where existing options are inadequate. All developed hardware was aimed to be low cost, whilst new software was designed with an emphasises on ease of use to practitioners, so as to encourage wider uptake. Finally, assessment of the workflow as well as all approaches used, was conducted with recommendations on future work provided.

The proposed workflow was represented by three aims, and subsequent chapters will each address one research aim. These aims are as follows.

1) Characterise the study site, and identify the factors likely affecting animal distribution on site.
2) Track animals on each site and test their respective assumed uniformity of grazing intensity.
3) Rank the environmental factors affecting the spatial distribution of animals on a given site.

Characterizing each site and identifying factors affecting animal distribution would allow for necessary baseline information to be collected about the environment likely affecting the animals. Testing the supposed uniformity of grazing intensity across a given site included exploring the dynamics of grazing behaviour at each site, and how successful current management prescriptions were likely to be. Ranking the environmental factors affecting the spatial distribution of animals on each site allowed a comprehensive understanding of what is driving the specific distribution of animals at each site. From this information, it was intended that possible recommendations on future management interventions be suggested. This would

lead to manipulation in the distribution and grazing behaviour of the animals in a way that better facilitates restoration according to specific site conservation goals. Specific chapters objectives required to achieve these overall aims are included in Table 1.2.

In order to best demonstrate the effectiveness of the workflow across different habitats, an upland study site with mixed habitat composition was assessed. Likewise, given the notable importance of peatlands to ecosystem services, another study site was an upland degraded Molinia dominated peatland site. This second site is indicative of many degraded peatlands within the UK, contains different conservation goals to the first site, and ultimately was the primary focus for developing the new workflow.

In deciding what animal should be used to test the workflow, it was agreed that sheep would be used as they had been present at both sites for many years and therefore their familiarity of the sites were established. At both sites, only the 'Welsh Mountain' breed were present and therefore this was the breed used. It should be noted that other breeds of sheep would have been acceptable as would other livestock species (e.g. cattle) to test the workflow, however as 'Welsh Mountain' sheep were present, this was what was used.

Table 1.2 Specific objectives for each experimental chapter

Chapter	Objectives
Chapter 2	Conduct initial descriptive surveys of each study. Determine the conservation goals of each site. Determine past and current management activities on each site. Explore, decide on, and use the most appropriate data collection approach for characterising the study sites in question. Explore the differences between the sites.
Chapter 3	Explore the hardware options for tracking animals. Develop or utilise tracking hardware that is low cost and capable of tacking large numbers of animals. Develop or utilise existing software which can process tracking data. Develop or utilise existing software which can be used by practitioners to infer grazing behaviour and characteristics (e.g. home ranging). Test the supposed uniformity of grazing intensity across a given site.
Chapter 4	Identify the most suitable modelling approach for the study objectives. Develop a conceptual framework for modelling. Prepare the data and select the most appropriate models. Evaluate model performance and the ecological realism of the results.

CHAPTER 2

SITE CHARACTERISATION AND VARIABLE PROFILING

2.1 Background context

The first step in understanding how best to manipulate animal behaviour together with resultant habitat restoration, is to understand in detail the characteristics of each individual site and collect baseline information. The importance of a baseline understanding of a site and related interactions with grazers on it is fundamental to accurately compare pre and post treatment, and therefore know the extent to which an intervention is working. As such, the first aim of the overall study (*Characterise the study site, and identify the factors likely affecting animal distribution on site*) will be considered here.

2.1.1 Why classifying environmental variables is important

Habitat selection is an expression response of animals to a large number of often interdependent variables which constitute their relevant environment (Shannon *et al.* 1975). These variables are typically behavioural or environmental in nature, and often interrelated (Peek *et al.* 1982). Variables may be biotic (e.g. vegetative cover, biomass of forage), or abiotic (e.g. slope, aspect, elevation); discrete categories (e.g. vegetation classes), or continuous arrays (e.g. distance to water, vegetation height) (Manly *et al.* 2007). Within the context of animal-environment interactions, these underlying habitat attributes are often referred to as 'resources' (Avgar *et al.* 2016). The specific usage of a resource is further defined as the quantity of a resource that is utilised by animal (or population of animals) in a fixed period of time, with resource selection occurring when resources are used disproportionately to availability (Erickson *et al.* 2001).

2.1.2 Why scale matters

When assessing the resources available to an animal, it is vital to consider the spatial and temporal scale an animal operates within. From a spatial perspective, whether an animal population is being studied at a regional, landscape, or field scale obviously has implications on the specification of resource availability (Manly *et al.* 2007). These scales may be set by: 1) the range capacity of a population; 2) physical boundaries-either natural (e.g. topography) or manmade (e.g. fences); and 3) through the study design (i.e. the scale most interesting biologically).

In general, the concept of resource selection assumes that different resources are not uniformly distributed (Orians & Wittenberger 1991). Therefore, as scale increases, so typically will habitat heterogeneity (Báldi 2008), which in most cases will then correspond

with more apparent resource selection occurring (at least to a certain size) (Boyce 2006). Conversely, at a field scale, the habitat will typically be more homogenous, and therefore selection will usually be less pronounced, and animal distribution more uniform as they maximise available resources. Temporally, how long an animal is present on a given site will often influence resource selection. This is because full exploration of an area takes time, and therefore new animals to a site will often make resource selection decisions on the basis of incomplete information (Orians & Wittenberger 1991). Also, the larger the area the longer it takes for an individual to explore. Animals which have been present in the same area for a while will often then have a competitive advantage as they are able to make more informed choices in their resource selection, both at a particular time, but also across seasons where resource availability may alter. This phenomenon is well observed and is termed home ranging (see Chapter 3 for further information). Though, the length of time it takes to become accustomed to a site is not easily discerned, as it depends on a variety of factors (e.g. species, site size, physiology, microclimate).

2.1.3 Important considerations for profiling the study sites

In attempting to specify resources and how they should be measured, it is important to consider the following points. Firstly, the scale of a site (i.e. the study area) must be determined. Whether the site boundary is physically imposed or imposed by the study design is also important. Secondly, whilst profiling a site, the characteristics of the animals' present must be considered. Though this is not a determinant of the resource availability, knowing their ages and respective history on the site can inform whether home ranging is likely to be occurring, and can help inform individual variation in resource selection should it occur. This also has relevance to the third point, which is that there must be an awareness of whether there are any current habitat management procedures implemented on a site. This may be related to the animals in the form of grazing prescriptions, or it could be other methods, which in turn may affect animal distribution depending on specific intervention. For example, vegetation removal through mowing or burning will often lead to new growth which could be exploited by the animals. Knowing this information is useful for understanding the attributable factors affecting resource use, which in turn can facilitate an investigation into the effectiveness of these interventions as tools for manipulating animal distribution. This is of course a key focus in this study. Fourth and finally, once each site has been thoroughly profiled, any pre-existing data that could be used to determine resource availability must be explored. Close consideration should be given to the resolution required in order to best

quantify the resources in accordance to the study scale. In the event that the current data does not meet the required resolution or is unsuitable in some other way (e. g. inaccurate), then other measurement techniques which can deliver the required data must be explored.

2.2 Existing data on sites

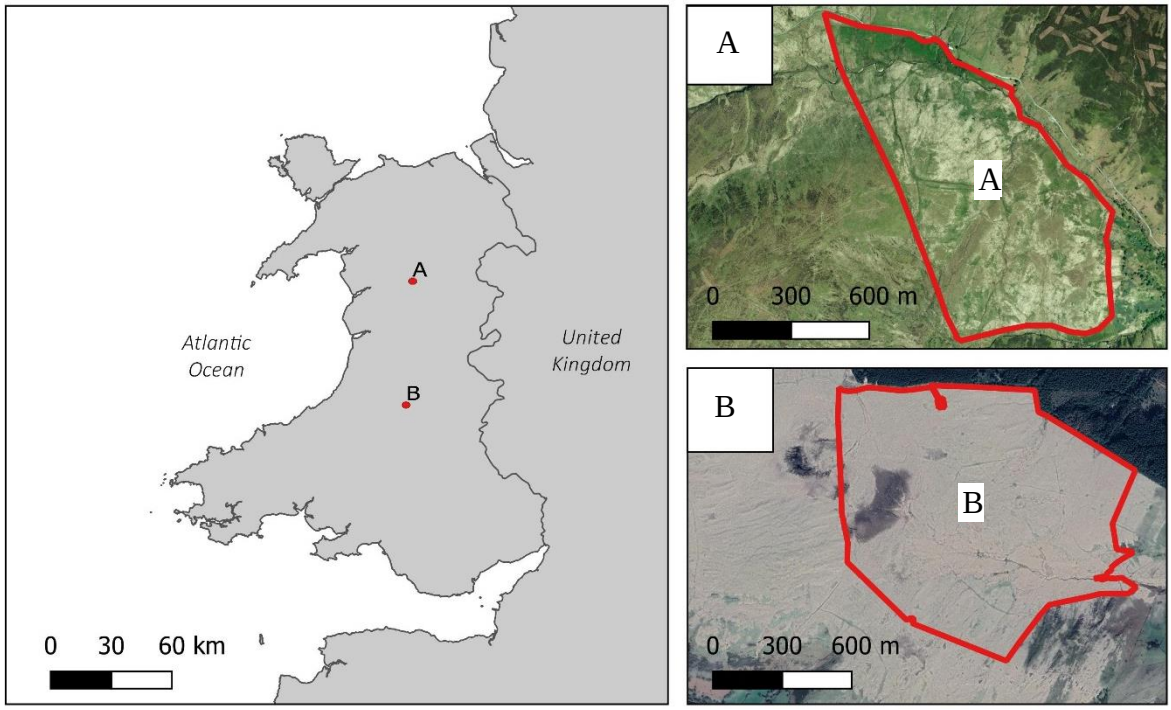


Figure 2.1 Site locations in Wales, United Kingdom. Red line in A & B = site boundaries. A = Ffridd Fawr, B = Penglaneinon. Site base layer = Google satellite.

2.2.1 Ffridd Fawr (Lake Vyrnwy)

2.2.1.1 Location

Ffridd Fawr is a 72 ha site near the Lake Vyrnwy reservoir in Mid Wales (52°47'49.34"N; 3°34'34.99"W). The elevation of the site varies between 400-509 m above mean sea level (AMSL). The site encompasses a section of valley bottom, a steep north-facing slope, and a shallow gradient plateau at the top. The whole site is enclosed by a post and wire fence.

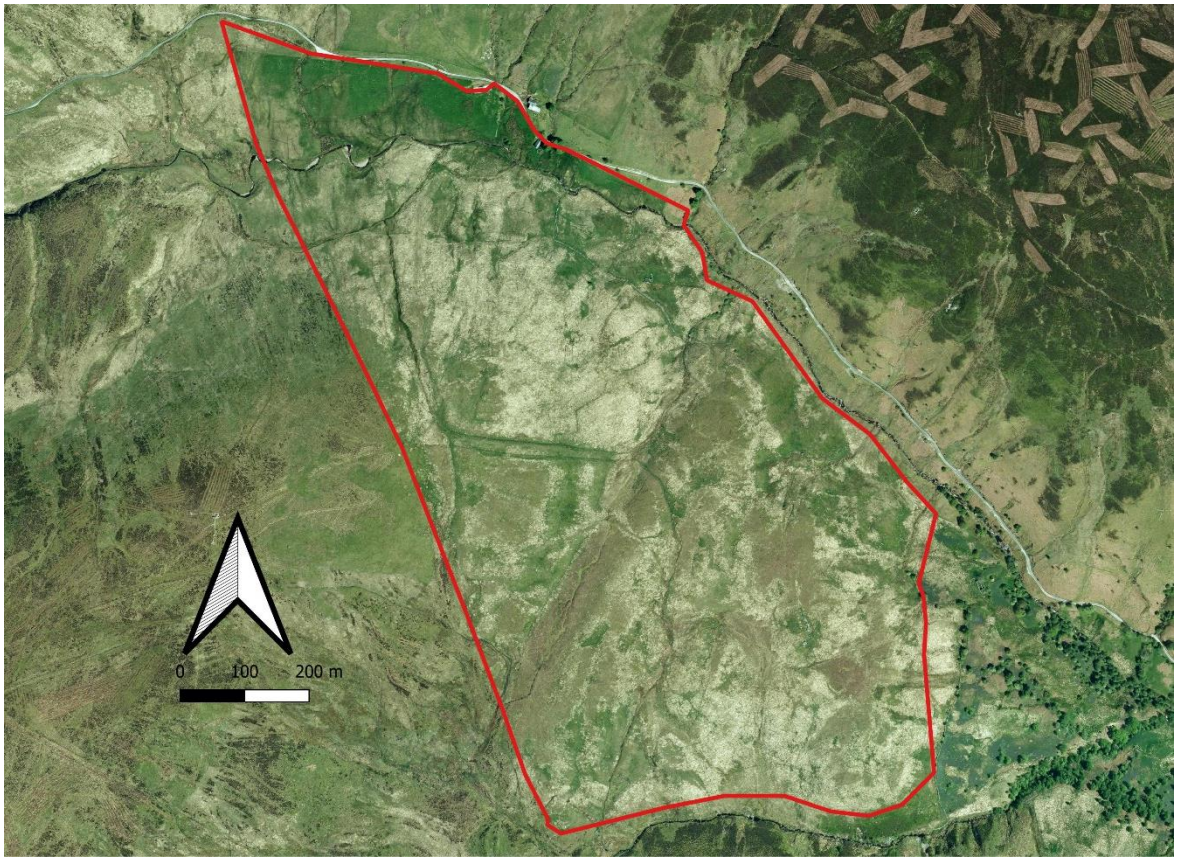


Figure 2.2 Ffridd Fawr site at Lake Vyrnwy. Red line = site boundary. Base layer = Google satellite.

2.2.1.2 Site description based on observations from initial visits

The vegetation at Ffridd Fawr is the more heterogeneous of the two sites, with a greater degree of patchiness occurring on site. The vegetation at the lowest section, along the valley bottom would be regarded as semi-improved acid grassland. This continues until the base of the slope, but then transitions into a more unimproved/marshy grassland as the slope gets higher. Along the top of the slope, sections of acid grassland persist, however there is also a noted presence of *Erica spp*. (bell heathers) and *Calluna vulgaris* (ling heather) with a general progression to a more wet heath habitat. Small sections of dry bog are also interspersed, particularly to the western side. Though there is *Molinia* on the site, it does not exist in detrimental quantities. *Pteridium aquilinum* (bracken) is present in patches, with most existing to the eastern side of site. With regards to above ground water present, there is a small river that runs along the bottom section of site, through the valley bottom. There is also a stream that intersects the entire site towards the top of the slope. Aside from this, there are a few small streams and puddles that are spread around the site; however their permanence is

unknown. There are also only a few trees present, with a small number of deciduous trees located along the edge of the stream on the eastern side of the site. Finally, a low number of manmade features exist on site. Most notably are two derelict stone buildings towards the east and north east of the site, with a stone wall running between them.

2.2.1.3 Site management

Ffridd Fawr is managed under the Welsh Government's 'Glastir' agri-environment scheme and grazing pressure is therefore applied in accordance with the LSU system. From the beginning of October to January there are no stock present on site given the weather during this time. From February to April, 80 barren ewes are present. These are replaced in the beginning of May by ~180 ewes with lambs at foot. These are on-site until shearing time in the middle of August. These sheep are then replaced with ~200 different ewes and lambs, which remain on site until weaning at the end of September. The prescribed stocking rates for the months the animals are present on site are 0.13 (minimum) to 0.26 (maximum) LSUs per ha. Welsh Mountain sheep are the only breed of sheep used on site. This grazing regime had been imposed for two years before this project commenced.

2.2.1.4 Conservation goals

Ffridd Fawr is managed by the RSPB, who manage a considerable portion of the larger Lake Vyrnwy estate. The two primary objectives for the RSPB at Lake Vyrnwy are: 1) the restoration of blanket bog habitats; and 2) the management of habitats to support key breeding birds such as *Turdus torquatus* (ring ouzels), *Numenius arquata* (curlews) and *Lagopus scotica* (red grouse).

The Ffridd Fawr site sits at the base of the main moorland parcels, and therefore is not part of the blanket bog restoration programme. It has however been identified as an area of interest for *N. arquata*. Directly adjacent to the eastern side of the plot is an area also managed for *T. torquatus*, however this is not strictly part of Ffridd Fawr. Specific objectives that relate to main moorland parcels as well as Ffridd Fawr are centred around creating a diverse sward structure to provide conditions for key species and prey species. This includes areas of short sward and invertebrate-rich fields which will benefit *T. torquatus*, and other areas with tall tussock vegetation that is open enough to allow birds to move around freely as suitable nesting habitat (such as *N. arquata*). Indeed, many of the recommended methods of creating suitable conditions for *N. arquata* were developed at Lake Vyrnwy (Fisher & Walker 2015). In their conclusions, these authors suggested a finer heterogeneous mosaic of short and long

vegetation to be optimal, though with consideration given to the scale of habitat being assessed. Where cutting/mowing is used to achieve this, they also advocated the use of rotational cutting instead of mowing existing short areas in order to alleviate the risk of disturbing birds that are already utilising these shorter areas. This understanding therefore presents an opportunity to develop a management tool whereby livestock on site can be manipulated to rotationally graze around the site to deliver this mixed sward height patchiness. No cutting/mowing was being undertaken at the time of the project start, nor had any happened within the recent years preceding.

2.2.2 Penglaneinon (Elan Valley)

2.2.2.1 Location

The Elan Valley is a 186 km^2 river valley situated within the Cambrian mountains in Mid Wales. Penglaneinon (literally translated as the fields on top of the riverbank) hill is a 112 ha site situated centrally within the estate, overlooking the Caban Coch reservoir (52°15'14.5"N 3°36'32.9"W). The site has an elevation of 380-440 m AMSL, with a predominately south-easterly aspect. Much of the site sits extremely exposed just below the apex of a low ridge situated to the southern edge of the site. Aside from this, most of the site has a consistently shallow gradient. The whole site is enclosed by a post and wire fence.

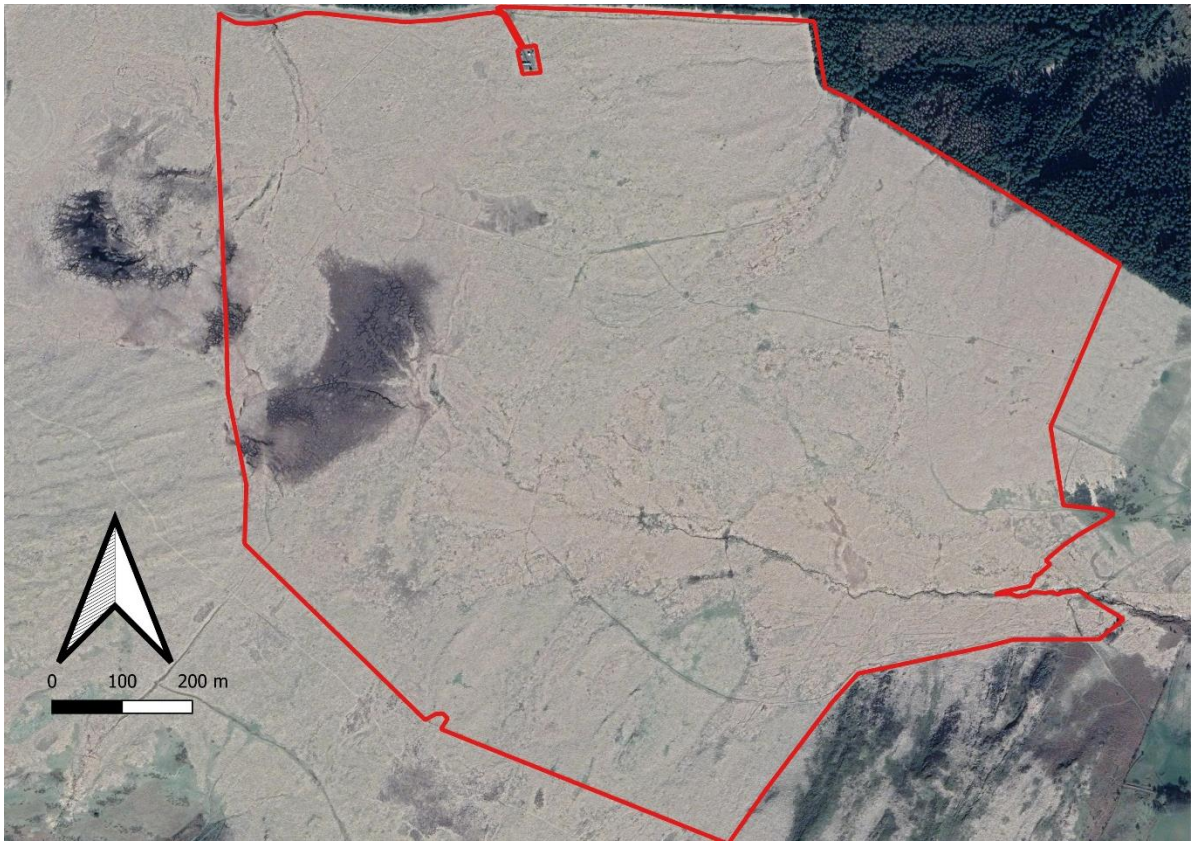


Figure 2.3 Penglaneinon site in the Elan Valley. Red line = site boundary. Base layer = Google Satellite.

2.2.2.2 Site description based on observations from initial visits

As a brief description, much of the vegetation on site could fall under the class of degraded marsh grassland since most of the site is almost singularly dominated by *Molinia caerulea* (purple moor grass; hereafter Molinia) (Figure 2.4), with small patches of *Deschampsia cespitosa* (tufted hair grass) (Figure 2.5), and *Juncus effusus* (soft rush) (Figure 2.4) interspersed within. There is however a sizeable section of blanket bog to the west of the site (See brown patch in Figure 2.3). *Trichophorum germanicum* (Deer grass) is present in most parts of the blanket bog, as well as different *Spagnum spp*. *Drosera rotundifolia* (round-leaved sundews) are also present in small quantities as are a number of other mosses, lichens and liverworts, which remain unidentified, and require future work to verify. The blanket bog also contains several tributaries that lead to a central stream which then runs the extent of site from east to west. Finally, though there are conifer plantations on the northern edge, these are not part of the site extent. There are small number deciduous trees present on site (Figure 2.4), however these are limited to singular sparse individuals which are most prominently situated in the gully of the stream which runs through the site. The site contains few

manmade features. However, there is a derelict radio station along the north edge of the site, which is enclosed by a fence, as well as a small network of 4x4 of tracks of varying sizes. As alluded to previously, almost all the identified above ground water is contained within the blanket bog tributaries and the stream running the length of the site. During early reconnaissance, an additional stream was identified to the north-west of the site, as well as a low number of other small water bodies (puddles), however it concluded that these were unlikely to permanent, and therefore not prudent to include as part of available resources.

Figure 2.4 Image of Penglaneinon site during winter months. Image taken from the southern side facing northwards. Yellow/white vegetation is *Molinia caerulea*. Small patch of *Juncus effusus* also present in front of tree.

Figure 2.5 Image of *Deschampsia cespitosa* present at Penglaneinon site. Image taken from southern side facing east during winter months.

2.2.2.3 Site management

Similar to Ffridd Fawr, the site is currently being managed under the 'Glastir advanced' agri-environment scheme. Grazing on the site is therefore prescribed under the LSU system, and has been operating under the same protocol for last ~5 years. Given the exposed and harsh climatic conditions on site during winter, there are no animals present on site between mid-October and March. All grazing is undertaken by Welsh Mountain sheep. From the start of April, ~fifty yearling (between 1-2 years old) sheep are present on site. These are joined by ~ 145 ewes with single lambs at the start of May. This stays the same until September, when the lambs are weaned (removed) from the ewes, and some of the yearlings will also be removed. Around mid-September to mid-October the number of sheep on site is usually ~150. The only time the sheep are handled during this period is for shearing ~mid-August, and for weaning ~mid-September. As per the LSU system, suckling lambs are not deemed to contribute to the number of grazing animals. The prescribed stocking rates for the months the animals are present on site are 0.13 (minimum) to 0.26 (maximum) LSUs per ha. The age of the ewes varies from 2-4 years.

2.2.2.4 Conservation goals

The valley is owned by the Dŵr Cymru (Welsh Water) company and is managed wholly by the Elan Valley Trust. A principal function of the estate is the provision of a clean water supply for Birmingham, and the whole area is managed in in order to protect the water quality. As a result, agricultural activities are low intensity, with an emphasis on management for environmental and recreation activities. As such, the whole area categorised within the now superseded Cambrian Mountains Environmentally Sensitive area (ESA), is a Special Protection Area (SPA) for wild birds, has Special Areas of Conservation (SAC) for habitats, and most of the land has been designated as various Sites of Special Scientific Interest (SSSI). The specific management objectives for Penglaneinon are predominately based around transitioning the marsh grassland into further blanket bog areas. The principal thinking being that this would result in greater water retention capability on the site, as well as improve the biodiversity present and carbon uptake potential. The best way of delivering this habitat transition is deemed to be the removal/reduction of *Molinia* across the site. It is reasoned that by doing this, the vegetative canopy will become more open, thereby allowing the previously barren understorey greater access to light, and therefore opportunity to grow. Through time this will allow key bog forming vegetation such as *Sphagnum spp.* to colonise (Figure 2.6).

Figure 2.6 New vegetative growth at Penglaneinon following removal of *Molinia* by mowing the previous year. New vegetation in the picture is predominately *Polytrichum commune* (common haircap moss).

Key efforts to reduce the *Molinia* has mainly been focused on manipulating the grazing intensity across the site to put more pressure on the *Molinia*, either through being grazed or through disturbance. However, the inability to increase stocking rates due to the LSU system has meant this has so far not been successful. The inclusion of mowing and vegetation removal in 2015 was undertaken in order try and facilitate new growth. It was hoped that by doing this, the sheep might then prevent *Molinia* regrowth by grazing the newer vegetative material. Though visual inspection of the cut areas 18 months after mowing indicate this to be the case (as seen in Figure 2.6), this is anecdotal evidence and remains to be quantified properly. If proven true, this would support the notion of further cutting on site, with aftermath grazing.

Another reason why being able to manipulate grazing is considered preferential by the land managers is their belief that the sheep are over grazing/or disturbing the blanket bog area to a detrimental level. Another management objective for them is to therefore be able to reduce

the distribution pressure on the blanket bog area by enticing the sheep elsewhere. Though it could be argued that such a result could be achieved by fencing the blanket bog section off from the main site, this would represent an expensive venture, and given the high soil moisture level of the site, need to be repeated fairly regularly. A query could be raised as to why a mixed species grazing approach has not yet been implemented on site (e.g. incorporate cattle and ponies) – given that many studies have extolled the benefits of utilising different livestock species together (e.g. cattle and sheep). It has been shown that different species select vegetation differently, and can even complement each other's preferences (Bakker 1989; Animut & Goetsch 2008; Fraser *et al.* 2014). For example, cattle often graze down long vegetation which sheep are then able to low graze (Putfarken *et al.* 2008). A predominant reason is that the enclosure fencing is not deemed adequately robust enough to withstand the pressure of larger herbivores. Therefore, in order to accommodate these animals, considerable resources would need to be invested in order to secure the site properly. Given the extent of the land within the estate, focusing such level of resources in one area is not something that can be undertaken lightly.

In summary, at 112 hectares the Penglaneinon site is but 0.16% of the entire Elan Valley estate yet it is representative of much of the landscape of the area. It is therefore believed that identifying the optimal approach for habitat restoration on this site would mean this method could then be applied to other areas of the estate.

2.2.3 Identifying the resources

Using this understanding of the overall features of both sites, how they are managed, and the respective landowner's objectives for them, it is possible to assess the factors/resources present that could potentially influence animal distribution. Most of these factors are common to both sites. There are after all fundamental resource requirements which animals require to live, as well as environmental factors which are prevalent on any such site. In brief, these are the factors and resources deemed important for consideration at both sites.

2.2.3.1 Above ground water

Water availability is fundamental to the survival of any living organism. With regards to mammals, this water is typically accessed through above-ground water bodies (as opposed to soil-based). Streams, rivers, ponds, lakes, or even large puddles can be included in this category. However, caution must be observed with any particularly small or shallow puddles as these may only be temporary in nature (e.g. when there is high rainfall), and instead the

focus should be on water bodies which are likely to be available for the majority of the time animals are present on site.

2.2.3.2 Vegetation class

What and why ungulates graze preferentially is a core question within ecology (Hanley 1982; Owen-Smith & Novellie 1982). Sheep in particular are noted for their high selectivity when grazing (Marteinsdóttir, Barrio & Jónsdóttir 2017). Assessing the different types of vegetation present on each site is therefore crucial. This could be in the form of habitat classes, or down to species level depending on the requirements of the study.

2.2.3.3 Vegetation greenness

Aside from the type of vegetation, it is also well documented that ungulates will typically preferentially select greener, more productive vegetation when they can (Bro-Jørgensen, Brown & Pettorelli 2008). Derived from a ratio of red and near-infrared reflectance (NIR), the normalised difference vegetation index (NDVI**)** is a well-established indicator of vegetation greenness, and related vegetation productivity (Pettorelli *et al.* 2005). It is therefore an obvious measure to include when assessing resource use.

2.2.3.4 Slope

Slope gradient is a recognised potential determinant of herbivore distribution in rugged terrain (Vavra & Ganskopp 1987). With an increase in slope gradient, there will typically be a greater energy demand during animal movement. However, some species will exploit higher gradient slopes for the advantages they can offer (e.g. visibility/protection against predators, particular vegetation present). Because of this, it is often reasoned that slope use is more related to associated abiotic or biotic features rather than the actual slope per se (Shannon *et al.* 1975). Nevertheless, it represents a valid factor that can affect animal distribution, whether it be direct or indirectly causative. Testing for correlation between the variables before analysis will inform us on the extent to which slope gradient is likely to be directly contributing.

2.2.3.5 Aspect

Like slope, aspect is another factor which is often associated with biotic and abiotic factors, particularly climatic conditions. Prevailing wind direction, as well as sun exposure will vary according to aspect, which causes can cause differences in the site microclimate. This will in turn then have effects on other variables (e.g. greater vegetation greenness may well occur on

south-facing slopes as a result of increased sun exposure). Though neither slope or aspect can be manipulated, measuring them can help interpret other factors such as vegetation greenness, or why vegetation class are distributed in the way they are on a particular site.

2.2.3.6 Shelter (Ffridd Fawr only)

Another factor inextricably linked with climatic conditions is shelter. The preferential use of shelter by animals to moderate adverse conditions on a site is well known (Stahl *et al.* 2002; Hirzel & Le Lay 2008). However, within this study it is deemed necessary to only measure shelter at the Ffridd Fawr site. This is because the *Molinia* at Penglaneinon is so dense and tall (>1 m in some places) that the sheep could feasibly shelter in any of this thick vegetation. It would therefore be difficult to accurately quantify all the shelter and would likely not be preferentially selected given the abundance of the *Molinia* on site.

2.2.3.7 Other variables

Practicality and available resources inevitably dictate how many and which variables can be measured. A key variable already mentioned which could have been included if resources were sufficient is local climatic conditions. Though not a variable that can be manipulated, understanding the precise role and influence of local weather conditions on sheep distribution would permit a much more complete picture as to how the animals are being influenced and how the biotic and abiotic variables link on site. The decision not to measure climatic conditions was not taken lightly. However, fundamentally there were no weather stations in the immediate vicinity to the sites, and therefore a considerable capital cost would have been incurred in installing monitoring stations. In future work, it is however a variable that should be considered. One such suggestion would be to strategically place a series of microclimate sensors around the site, which through interpolation between them, would allow for climatic gradients across the sites to be assessed. Another variable which could be also considered in future is vegetation height/density. Observing whether the animals would be influenced by this would be useful, however more importantly for both sites, being able to monitor sward structure would provide clear indications of how manipulations/other habitat restoration methods were working. At Penglaneinon, this would include seeing whether *Molinia* patches were reducing in height and density, whilst at Ffridd Fawr, the vegetation structure and density could be monitored to actively inform the management of both *N. arquata* and *T. torquatus* habitat areas.

2.2.4 Measuring the variables

Though the use of quadrat-based surveying methods has long been utilised to determine vegetation composition and condition, they lack resolution on large sites - as practical limitations often result in them being minimally representative. Given the size of the sites and the variables to be measured, the use of raster-based data is the ideal form for the scale of the sites. Simply put, a raster consists of a matrix of cells (or pixels) organised into columns and rows (or a grid). Contained within each cell is a value representing information, which can be either categorical (e.g. vegetation type), or continuous (e.g. distance to water). Examples of raster data sources include digital aerial photographs, imagery from satellites, or even scanned maps. When incorporated into Geographical Information Systems (GIS) software, this data can be georeferenced (so each cell is spatially referenced), then analysed, managed or manipulated to highlight or infer aspects of geographical components within the data. Raster's therefore represent an ideal way to visualise, interpret, and validate the variables that require studying.

2.3 Phase 1 habitat surveys

With geospatial data, there is typically a compromise between the resolution of data required, and the cost incurred to acquire it. Therefore, a prudent first step is to explore whether any freely available data already exists that could be used for analysis. This may be data collected directly on the sites or accessed through freely available sources. Any data gleaned should then be evaluated on its ability to identify the necessary variables, and the resolution assessed to determine if it is sufficient considering the size of the sites, and the objectives to be met.

2.3.1 Original Phase 1 habitat surveys

The Phase 1 habitat classification is a standardised system of recording semi-natural habitats, developed by the Nature Conservancy Council. It was designed to cover large areas of countryside relatively rapidly, and it presents the user with a basic evaluation of habitat type; thereby providing an indication as to the potential importance of a given area for nature conservation (JNCC, terrestrial habitat classification schemes). For Wales, it represents one of the first comprehensive assessments completed of habitat type nationally. Between 1987 and 1997 over 80% of the land surface was surveyed in the field using Phase 1 methodologies (Stevens *et al.* 2004). Given this level of national coverage, and that the fact the data is freely

available to download (https://lle.gov.wales/home) it was deemed a prudent first step to explore these maps.

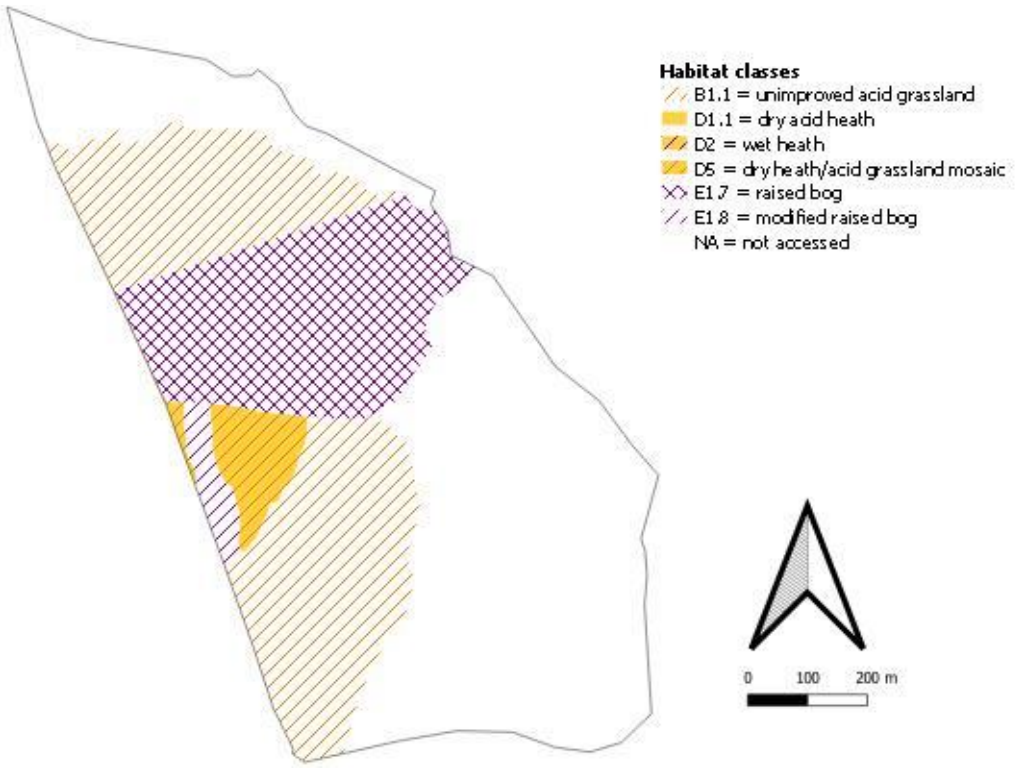


Figure 2.7 Terrestrial Phase 1 habitat Survey for Ffridd Fawr. Data accessed through Lle geo-portal (https://lle.gov.wales/home). Habitats were styled according to Natural Resources Wales (NRW) Phase 1 standard.

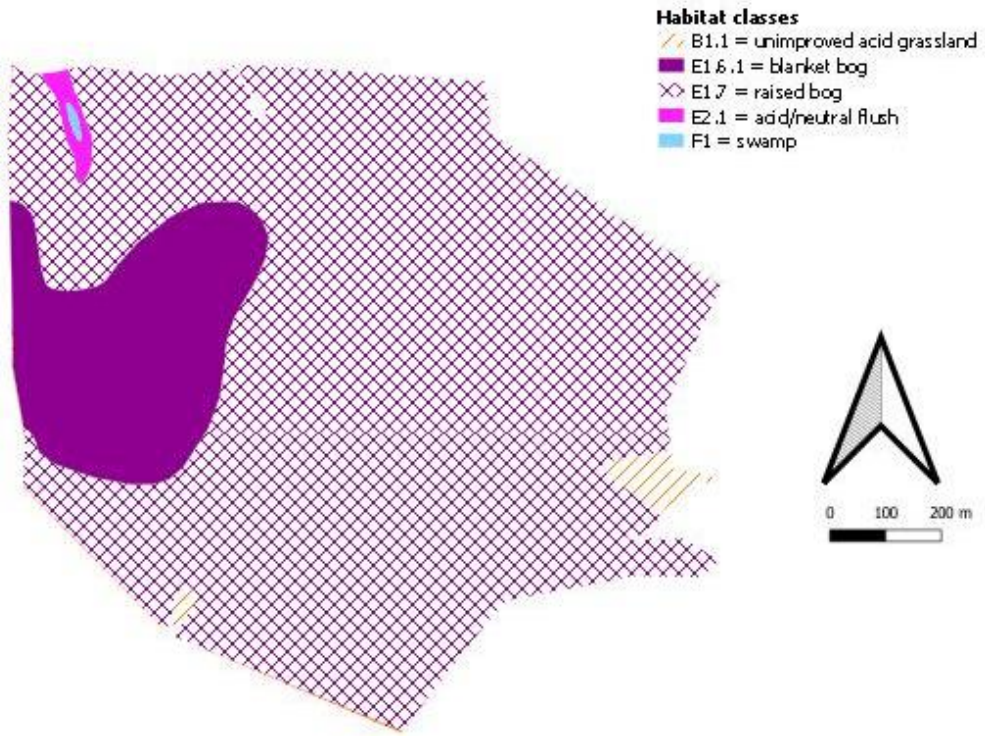


Figure 2.8 Terrestrial Phase 1 habitat Survey for Penglaneinon. Data accessed through Lle geo-portal (https://lle.gov.wales/home). Habitats were styled according to Natural Resources Wales (NRW) Phase 1 standard.

As is evident from Figure 2.7, the main limitation making this data unsuitable is the fact that not all of the Ffridd Fawr site is covered by the survey (i.e. not accessed), and therefore is incomplete. Aside from that, several other concerns exist. The first of these is that many of the stated habitat classes do not reflect the current state of the site. This is most apparent in Figure 2.8, where the most prevalent class is listed as being raised bog. Though achieving this habitat type is a key management objective for this site, in its current state it would more likely be classed as 'marshy grassland-*Molinia* dominated (B.5.2)', given the level of *Molinia* present. Aside from this, there is also the problem that some classes are missing within each site survey, one such absentee in both sites is standing water. In the case of the water bodies (e.g. streams/rivers), the issue of not recognising the classes is likely due to resolution. However, in the case of the vegetation habitat classes (such as the misidentification of marshy /*Molinia* dominated grassland as raised bog at Penglaneinon) this could well be due to the age of the surveys (Phase 1 surveys in Wales were completed in 1997), and therefore the habitats may well have changed during this time.

2.3.2 Updated Phase 1 habitat survey

The updated Phase 1 map of Wales was initiated in part to solve these concerns. In their study, Lucas *et al* (2011) performed an object-orientated rule-based classification of multi-temporal satellite sensor data. This classification was then fused with existing ecological knowledge of vegetation distributions (e.g. the original Phase 1 survey data) through using a combination of thresholds, Boolean operations, and fuzzy membership functions. The reported result of this process was a higher resolution (10 m spatial resolution), more up to date (satellite data taken from 2003 - 2006) national habitat map, that was indicatively estimated to be >80% accurate overall. Given these metrics, it seemed a sensible option to explore the potential of these data.

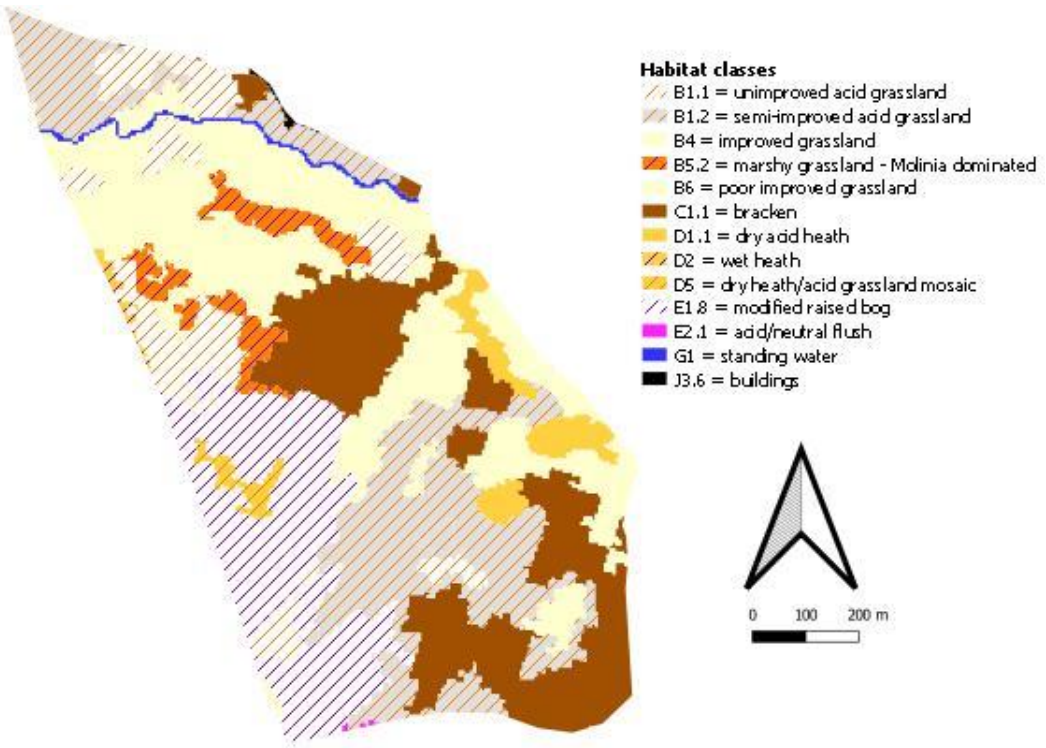


Figure 2.9 Updated Phase 1 habitat map for the Ffridd Fawr site. Data obtained from Lucas *et al* (2011). Habitats were styled according to Natural Resources Wales (NRW) phase 1 standard.

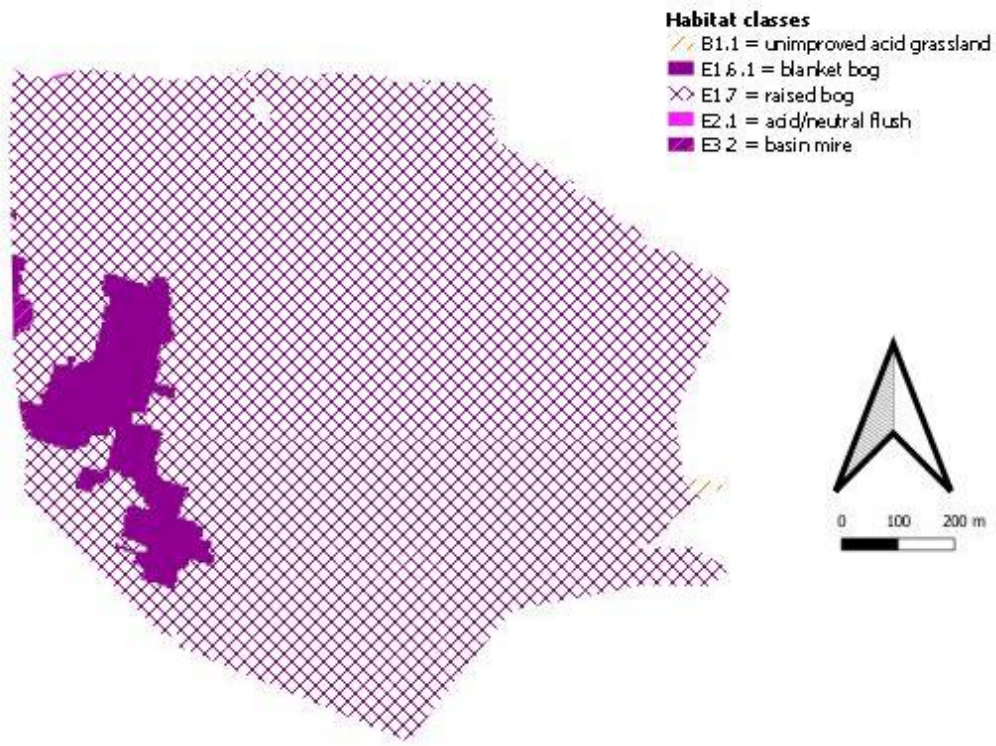


Figure 2.10 Updated Phase 1 habitat map for the Penglaneinon site. Data obtained from Lucas *et al* (2011). Habitats were styled according to Natural Resources Wales (NRW) phase 1 standard.

At Ffridd Fawr, there is a notable increase in the number of habitat classes presented in the updated Phase 1 (13) compared to the original Phase 1 (6 classes). Though this can be partially explained by the fact the north and eastern sides of original Phase 1 were not classed in the original Phase 1, the degree of patchiness is visibly greater in the updated Phase 1 then the areas surveyed in the original Phase 1. There also does not appear to be much spatial overlap for habitat classes between the two surveys at Ffridd Fawr. Additionally, though standing water is now included in the survey it does not include all standing water seen on site, most notably the stream intersecting the site higher up. Finally, the distinction between certain classes in some cases is imperceptible (e.g. between improved grassland and poor improved grassland), and therefore unhelpful in interpretation.

Conversely, the updated Phase 1 map for Penglaneinon (Figure 2.10) is similar to the original survey (Figure 2.9) in many ways. The core areas and size of the two main habitat classes are broadly in agreement, though the blanket bog profile does differ between the two. Furthermore, the updated Phase 1 also classes the main habitat present as ‘raised bog’ and not ‘marshy grassland - *Molinia* dominated’, whilst similarly failing to recognise any standing

water on site. The only main difference between the habitat classifications of the two surveys is the removal of the small acid/neutral flush, and swamp patch in the updated Phase 1 map.

Though the level of detail and definition appears greater in the updated Phase 1, both survey methods lacked the necessary resolution required for this study. Furthermore, the contradictions in the classifications and the omission of key classes known to be present creates uncertainty in the validity of these data. With all this considered, it was clear that neither Phase 1 data sets was adequate for either study site. Finally, though data from more intensive field surveying (e.g. Phase 2 or 3) may be sufficient, there are none known to have been collected at either site.

2.4 Other data sources

When all the variables to be measured are considered, it also clear that the Phase 1 data could at best be used to measure one variable (vegetation class). With regards to the other variables, alternative approaches would need to be utilised. It is feasible that standing water, and shelter (on Ffridd Fawr) could be identified manually, GNSS logged then digitised into a raster format, however this would be laborious and given the extent of the site, may still result in omissions. Slope and aspect are typically measured using digital elevation models (DEM). Given their free accessibility and global coverage, the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) and Shuttle Radar Topography Mission (SRTM) databases are the most widely used digital elevation models (Szabó, Singh & Szabó 2015). However, their respective resolutions are only ~30 m and 90 m. Given the scale of the study sites, and the knowledge that the accuracy of aspect and slope data decrease with lower DEM resolutions (Chang & Tsai 1991), these databases are not ideal in these circumstances. Wales does have a freely available Light Detection and Ranging (LiDAR) composite dataset (https://lle.gov.wales/home) which can be used to estimate slope and aspect. The resolution for data is high, with some areas measured to a 25 cm level. However, the coverage is patchy and neither sites are within the surveyed area. Finally, to measure vegetation greenness, NIR and red band spectral data is required. Given the seasonal variation within productivity, this variable would need to be measured during the study period. From a freely available perspective, the only realistic option would be that of free satellite data, such as that from Sentinel-2. However, like the SPOT-5 data used in the updated Phase 1 study (Lucas *et al.* 2011), the resolution of Sentinel-2 is only 10 m. Given the concern over the spatial resolution

of the updated Phase 1 data, it would be inappropriate to use the Sentinel-2 data, either in calculating NDVI or indeed performing an updated vegetation classification for the sites.

2.4.1 Commercially available data

Utilising these freely available data sources is clearly not enough to measure all the identified variables effectively and accurately. As such, the opportunities commercial approaches can deliver needed to be explored, with consideration given to the balance between the advantages of each an approach and the cost incurred to deliver the outputs. In order to measure all variables identified, multispectral data (data within specific wavelength ranges across the electromagnetic spectrum) covering both sites is essentially required (especially for NDVI as that requires red and NIR bands), as well as a DEM for measuring the topological features (i.e. slope, aspect). The resolution of these data is key, with realistically a <1 m pixel sampling size required, particularly for measuring the above ground water bodies. Regarding the former, there are several satellite options that satisfy this resolution requirement within the private domain. Digitalglobe's GeoEye-1, Worldview-2 and Worldwiew-3 satellites all deliver enough ground resolution for the requirements of this study with ground sampling distances (GSD) for multispectral images at 1.65 m, 1.84 m, and 1.24 m respectively. However, at the time of writing the cost per single image acquisition of each site is expensive. This is primarily due a minimum order size of 25 km^2 being set. As a result, though prices per km^2 are relatively low (~£13.40 - £17.25) for these satellite options, per acquisition this equates to ~£335 - £430 per site (based on 2020 costings). It should be stressed that this is based on the order of archival data (i.e. taken at least 90 days previous to the order date). Recent image collection (<90 days) costs more, with prices ~£23 per km^2 common (minimum order = £575). Further to this, satellite imagery is affected by cloud coverage, and therefore partial or complete obscuration of land area on desired acquisition dates is not uncommon, particularly for Wales where overcast conditions are frequent. Image acquisitions during low or no cloud coverage can be requested, however these will limit the dates available, and often incur extra cost. High resolution DEMs are also available from commercial satellite operators. However, as with the purchase of imagery, minimum order sizes are enforced, and therefore the cost per acquisition is again high. One such example is the Vricon digital surface model (DSM) and digital terrain model (DTM). It has a sufficient resolution of 5 m (horizontal/vertical accuracy = < 3 m Linear Error 90%), but the cost is ~£11.49 per km^2 with a minimum order size of 100 km^2 equating to a cost of ~£1150 per site.

2.4.2 Data collected using unmanned aerial vehicles (UAVs)

An increasingly popular approach to collecting the type of data required is the use of unmanned aerial vehicles (UAVs). As sensor platforms, UAVs can support a multitude of sensor types and achieve image resolutions well below 1 m GSD. Data can be collected as frequent as the user wishes, as long as weather conditions permit, whilst the imagery derived does not suffer the issues of cloud obscuration that satellite imagery does. While cloud coverage can affect local lighting conditions within images, hardware implementations such as down-welling light sensors (DLS) are used to compensate for this by measuring the incoming light for each image. Aside from providing high resolution imagery for measuring variables (e.g. vegetation class), high resolution DEMs can also be constructed from the imagery through the photogrammetric process of structure from motion (SfM). SfM is a technique for estimating three-dimensional structures from two-dimensional image sequences. It is typically completed by matching features within multiple images, thereby facilitating a reconstruction of the 3D-geometry within a scene (Mancini *et al.* 2013). The cost of using UAVs for such purposes varies enormously depending on the design of the airframe and specific integrated sensor required. However, the increasing range of designs available, both commercially and from a self-build capacity, means users can implement solutions which best suit the needs of their data requirements and budget.

With all this considered, it was apparent that exploring a UAV based approach to collecting the required data on both sites represented the most suitable measurement option. As a result, a low-cost UAV solution and sensor integration for collecting imagery on both sites was devised and deployed at both sites. The potential of the data collected to provide suitable raster outputs for each identified variable was then assessed, both using the UAV data alone, and in combination with other sampling methods. These environmental layers are presented within this chapter, and the merits of the approach taken, plus recommendations for future repeat sampling discussed.

2.5 Unmanned aerial vehicle design

2.5.1 Choosing the most suitable UAV design

In deciding what UAV would be most suitable for this study, there were several considerations:

1) **Survey area size**. Given both sites were in excess of >70 ha, it was clear that a fixed wing UAV design would be best suited. UAV aircraft can be assigned into either two categories: fixed-wing or rotary wing. Fixed-wing's resemble a typical airplane; that is, they have a rigid wing design and fly by generating lift from the UAVs forward airspeed. This airspeed is generated by forward thrust from a turning propeller and is adjusted using control surfaces built into the wing itself (e.g. ailerons, elevator and/or rudder). A rotary wing on the other hand features rotor blades that revolve around a fixed mast, known as a rotor. The number of rotors varies between different designs, but one (helicopter) or four (quadcopter) rotor setups are common. Rotary and fixed wing designs both have their strengths and weakness, but the main reason why fixed-wings are more suitable for surveying larger areas is because of their greater range capability. Because fixed-wing's require less power to stay in flight and operate at a higher airspeed, they can typically fly much longer and therefore further than rotary wings. Though dependent on the survey characteristics (e.g. operational altitude, image overlap etc.) a fixed-wing may well be able to survey >30 ha from a single battery use, whereas a standard quadcopter will usually only be able to complete <5 ha.

2) **Camera type**. The second consideration is the type of camera to be mounted on the UAV and therefore the payload capacity the UAV would require. Whether this camera would attach into the UAV power supply or require its own dedicated battery would also influence the weight. It was reasoned that a 500 g payload capacity would be enough to support a multispectral camera required for this study.

3) **Terrain following**. Thirdly, it was understood that the steep sided topology at the Ffridd Fawr site would require the UAV to have terrain following capability. This was to minimise the chances of UAV collision, but also to ensure that an even above ground altitude was maintained throughout the survey. This is crucial as merging images into a stitched orthomosaic requires enough overlap between the images (which would alter if the altitude between images varied).

4) **Weather variability**. A fourth consideration was that of the typical weather conditions on site. Given the exposed nature of the sites, unfavourable/windy conditions were common, and it was decided that the UAV should have a robust enough design to handle challenging airspeeds. It was determined that a Delta wing design would be most suitable. Shaped in the form of a triangle, this design has unique aerodynamic characteristics and structural

advantages for operating in high winds. Chiefly, the rearward sweep angle lowers the airspeed normal to the leading edge of the wing, thereby giving it stability at speed.

5) **Cost.** Finally, given there was now a narrowed specification for UAV, a decision was required as to whether a commercial product, or self-build solution be utilised. This was ultimately determined by cost. The price of commercial fixed-wing UAVs varies considerably, but at the time the study was conducted they were in a range of £8000-£60,000 for a design that matched the required specification. Conversely, it was estimated that a self-build option could be completed for <£1000. Given the budget constraints of this study, but also the awareness that many other comparable projects operate under similar cost restrictions, it was concluded that exploring a self-build option could provide insight as to whether a low-cost solution would be sufficiently reliable and provide adequate data quality for subsequent analyses. This would act as an indicator to other practitioners as to the viability of low-cost fixed wing solutions, who previously may have been unable to afford commercial alternatives.

2.5.2 Design and construction of the UAV

The UAV concept adopted was first documented in Ryan *et al* (2015), with the current iteration featuring many updates in both software and hardware compared to this original design. Modifications were also made to make the UAV more suitable for the terrain and environment in which the sites were located.

2.5.2.1 Airframe assembly

The UAV airframe was a publicly available Skywalker X8 (Skywalker; Hubei, P.R. China). Made from expanded polypropylene (EPP) foam, it has a delta wing shape, and a wingspan of 2.12 m. The airframe came dissembled and was glued together using IMPACT adhesive (Bostik Ltd, Stafford, United Kingdom), Gorilla Glue expanding foam (Gorilla Glue, Ohio USA), and epoxy resin. The Skywalker X8 features two carbon tubes to provide additional structural integrity to the aircraft, which span the fuselage and main body of both wings. However, several additional reinforcement modifications were also made. Wing-fuselage joints, motor housing, and servo arm bases were all constructed out of 3 mm plywood (instead of supplied plastic components). Carbon tubing also protected the servo rods. Additional strengthening was achieved by covering the entire aircraft in self-adhesive glass fibre reinforced cross-weave tape. An additional layer of self-adhesive black polyester tape

(on the lower half of the aircraft) and orange self-adhesive polyester tape (on the upper half) were added to promote visibility of the aircraft during flight, and to prevent UV degradation of the cross-weave tape. Finally, all electronic components within the UAV were secured in place using hook-and-loop fasteners (often referred to by the genericised trademark 'Velcro').

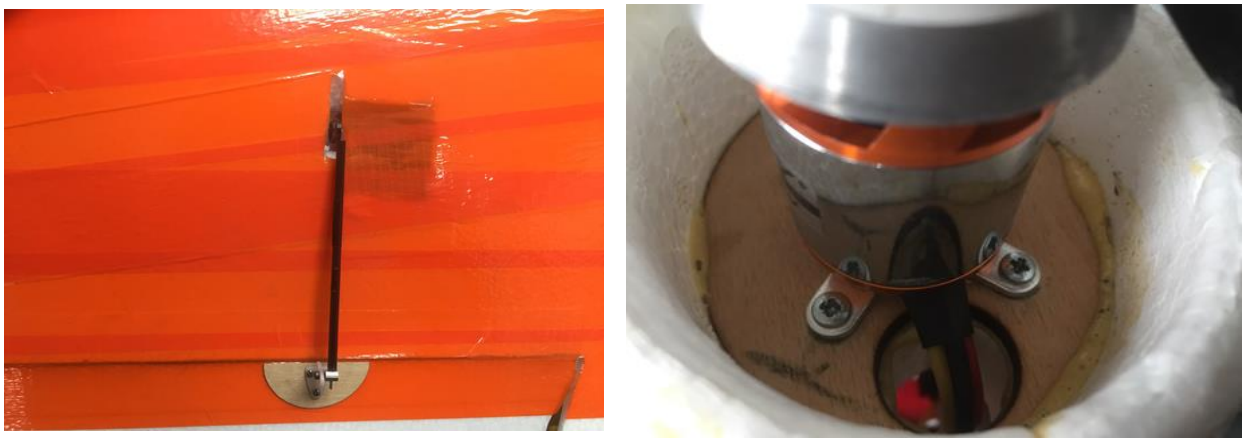


Figure 2.11 Reinforcements to Skywalker X8. On left image, note the plywood servo arm mount, and carbon tubing lining the servo rod. On right image - note the plywood motor mount.

2.5.2.2 Powertrain configuration

The powertrain (main components that generate power and deliver that power for propulsion) for the UAV included a 14.8v 14000 MAh (milliamp hour) Li-Ion battery, a 60A brushless speed controller, a 1100 kv (910 watt) motor, and a 28 x 18cm folding propeller. XT60 connectors were used between the battery and electronic components. The combination of the high revolution motor and small propeller was chosen as a compromise between the desire for sufficient thrust on launch for novice operators and challenging conditions (e.g. changing wind direction), whilst attempting to maximise efficiency in flight (by having less drag from the smaller propeller). Under this configuration, a maximum flight time of ~ 1 hr was achievable, depending on weather conditions and mission characteristics. These factors also influenced flight speed, with the average cruise speed being ~18-25 m/s. The minimum stall speed was 12 m/s.

2.5.2.3 Autonomous flight setup

Autonomous flight capability was used to provide flight stabilisation; altitude control; and GNSS navigation. The autopilot software used was the open-source project; Ardupilot (http://ardupilot.com/), operating on a Pixhawk 1 (https://pixhawk.org/) hardware module.

Flight stabilisation was managed using a triple axis accelerometer and gyro within the inertial measurement unit (IMU) of the Pixhawk, whilst a GNSS sensor and barometric pressure sensor were used for altitude control. Ardupilot features a dual-level proportional-integral-derivative (PID) controller architecture. The higher-level controls navigation, whilst the lower levels controls stabilisation. The radio-control (RC) interface consisted of a Spektrum dx6e controller bound to an AR610 receiver (Horizon Hobby, Illinois, USA). Two multi position switch outputs were mixed to provide six discrete values on a single RC channel to control the autopilot flight mode behaviour. These included; manual flying; lower level assisted flying (stabilise); higher level assisted flying (fly-by-wire-'A' (FBWA)); a fully autonomous mode (for waypoint following in surveys); return-to-home (RTH); and an autotune function used for PID parameter tuning. It was necessary to tune the PID parameters in initial flight testing, so they were optimised to the dynamics and mass of the airframe. This was to promote consistent stabilisation of the aircraft by minimising flightpath weaving (higher-level controller) and pitch/roll oscillation (lower-level controller).

2.5.2.4 Camera integration and survey setup

The autonomous flight mode allows pre-programmed 3D way-pointed mission plans to be executed on the UAV. Typical UAV survey plans consist of a series of stacked lines which aim to ensure complete coverage of a site when taking imagery. Distance between lines is influenced by the camera sensor specification (i.e. image size, focal length), and the degree of overlap required between images for accurate stitching to occur. The camera utilised within the study was a RedEdge multispectral camera (MicaSense; Seattle, USA), which was mounted in the airframe, facing nadir (90° to the ground) through a hole cut into the fuselage. The RedEdge is a multispectral imager with five sensors at 1.6-megapixel resolution; red (R); green (G); blue (B); near-infrared (NIR) and red-edge (RE). The integration featured an attachable GNSS sensor to enable image geotagging, as well as a combined magnetometer/downward light sensor (DLS) (MicaSense; Seattle, USA) mounted to the top of the fuselage to record angle during image capture and solar irradiance throughout the image runs. Combined with image captures of a supplied calibrated reflectance panel (CRP) before each image run, this provided necessary data for atmospheric correction. The camera was powered by the main UAV Li-Io battery, however a universal battery eliminator circuit (UBEC) was used to provide a regulated voltage supply from the flight battery to the camera. This was included because of initial testing which found the camera peak draw when taking

images to be enough to disrupt the autopilot system - particularly, when the autopilot regulator was used to supply camera power.

The standard level of side and frontal overlap for image stitching is 75%, so the survey lines for each site were therefore spaced accordingly to ensure sufficient side overlap. For frontal overlap, the capture trigger mode on the camera was set to 'overlap' which causes the camera to automatically take captures within 50 m of target altitude and calculates distance between each capture to ensure 75 % frontal overlap. Finalised missions were uploaded onto the aircraft from a laptop via a radio frequency telemetry micro air vehicle link (MAVlink). This telemetry link also allowed system status and mission progress to be followed in real time during surveys (Figure 2.15).

2.5.3 Regulations for UAV operations

A permission for commercial operations (PfCO) authorisation was not required from the Civil Aviation Authority (CAA) at the time of surveying. Local land permission was sought and granted from the RSPB and Natural Resources Wales (NRW) for the Ffridd Fawr site, and from the Elan Valley Trust for Penglaneinon. As the Ffridd Fawr site was within a SSSI (site of special scientific interest), the Berwyn SAC (special area of conservation), and SPA (special protection area), it was necessary to have the permission of NRW under section 28E of the Wildlife and Countryside act (1981), and in accordance with regulation 24 of the conservation of habitats and species regulations (2017). Given the frequent presence of low-flying military aircraft in both areas, notice of surveys at both sites were logged prior with the Ministry of Defence (MOD) low flying cell division. The CAA's standard drone code for all UAV operations was observed for the site surveys. Chiefly, the UAV flight height was set 120 m above ground level (AGL) and did not exceed this, whilst mission plans were also designed so that the 500 m UAV distance from the user was not exceeded. Given the size of both sites, this meant splitting the sites into two during flight planning, so each site required two flights to complete (Figure 2.13 & Figure 2.14). The topology of Ffridd Fawr (Figure 2.12) was a concern for one of the surveys in particular, as there was a ~100 m variation in altitude across the mission. From a regulatory perspective (to ensure 120 m AGL was not exceeded), and to ensure sufficient image overlap, it was necessary to employ terrain following during the surveys. Fortunately, Ardupilot contains a terrain following parameter which mission planner utilises through the global SRTM database for terrain. The global grid spacing of this data is 3 arc-seconds (around 100 m).

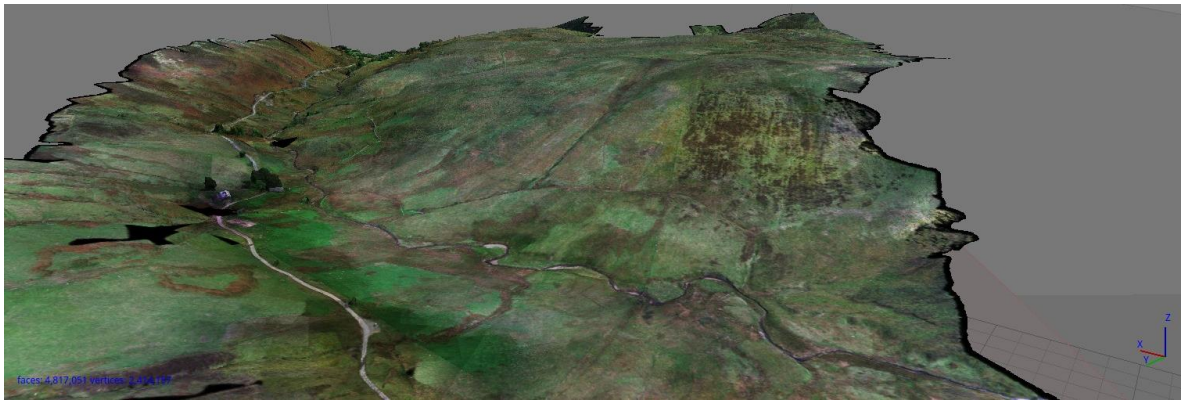


Figure 2.12 3D model of the Ffridd Fawr site and surrounding area produced from Ffridd Fawr imagery in Agisoft Photoscan (Agisoft LLC, St. Petersburg, Russia).

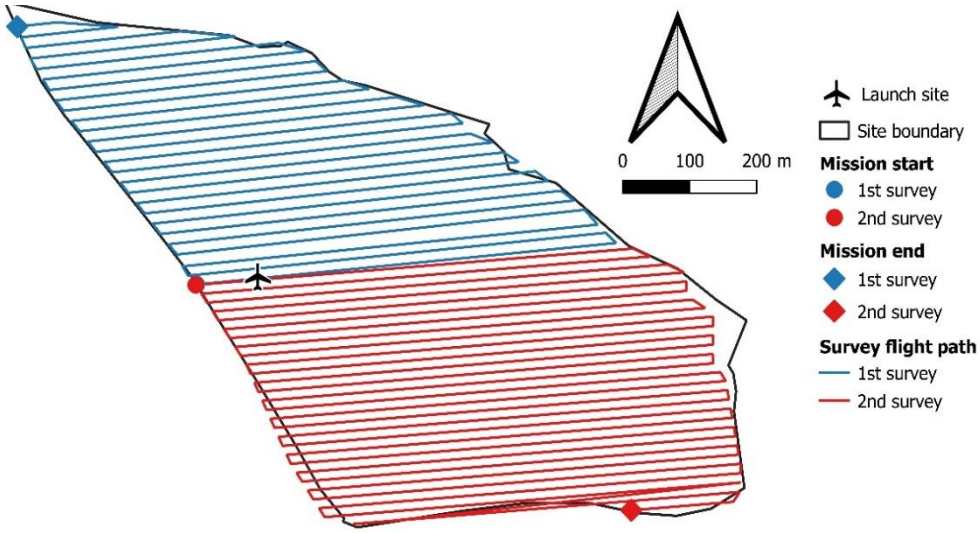


Figure 2.13 UAV Survey information for Ffridd Fawr. Note that 1st and 2nd survey starts are located in the same place.

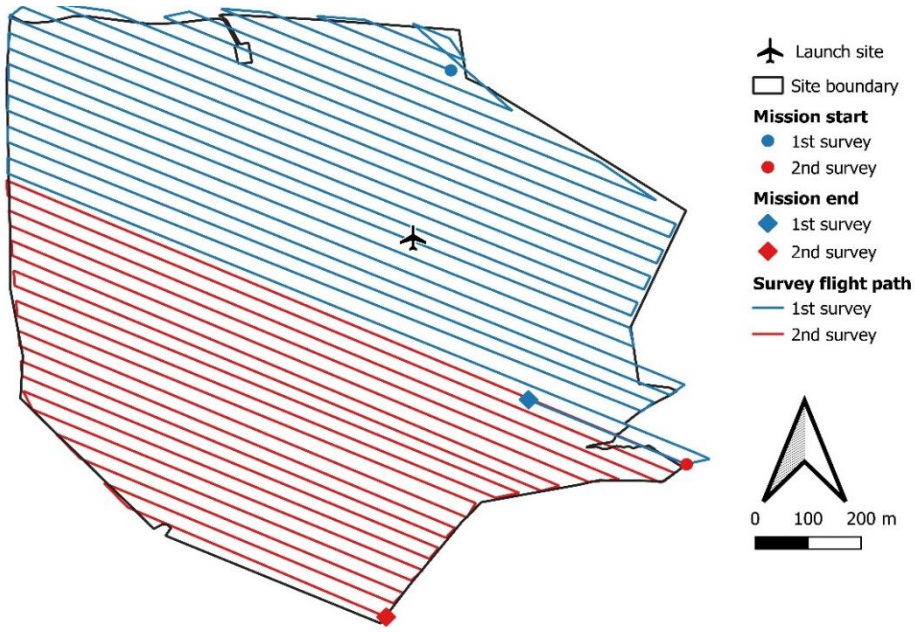


Figure 2.14 UAV Survey information for Penglaneinon

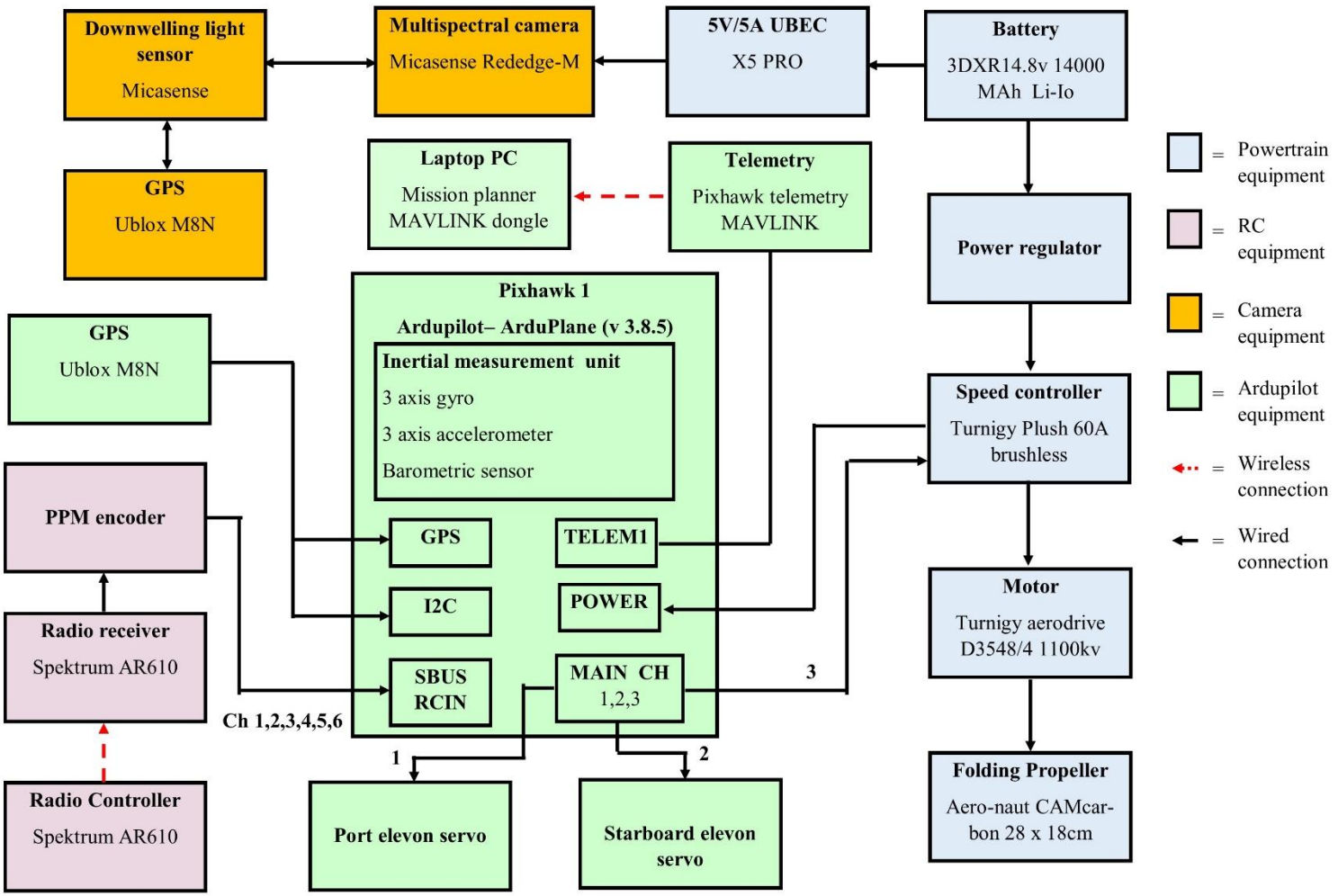


Figure 2.15 Logic diagram of the bespoke Skywalker X8 electronics

2.5.4 Procedure for UAV operation during surveys

During all missions, the UAV was launched and landed in FBWA mode. This allowed the pilot control of the UAV during the critical parts of the flight, whilst the autopilot provided enough stabilisation support to ensure a smooth transition. The UAV was hand launched by an assistant as it allowed the pilot to focus fully on operating the flight controller. Launch direction was always facing the prevailing wind, and flights were undertaken so long as wind speeds did not exceed 20 mph. Once the UAV was successfully launched and had reached survey altitude (120 m), the controller was switched to autonomous mode which caused the UAV to begin surveying. Given the size of the survey grids, it was necessary for the pilot to follow the progress of the UAV below to ensure line of sight was not exceeded. The assistant meanwhile monitored the telemetry feed on the laptop to monitor system status and survey progress. On completion of the survey, the UAV was belly landed. The steady characteristics of the FBWA mode ensured the UAV did not stall when flight speed was reduced, whilst the vegetation on ground ensured the underside of the airframe was not damaged on contact.

2.5.5 Initial testing and surveying

Due to an extremely wet summer, opportunities for surveying in the first year of the study (summer 2017) were restricted. Further to this, though two missions were conducted at Penglaneinon, flaws in the assembly (e.g. motors not secured sufficiently, Ardupilot not configured correctly) became apparent. As a result, no imagery was collected. Over the following winter a new UAV was constructed. Modifications were made, such as the plywood reinforcements, and the Ardupilot configuration was closely assessed. After initial testing of the new UAV at Penglaneinon proved successful in April 2018, it was decided that the multispectral camera be included in the following missions. The first two surveys with the camera however were not successful. In the first survey, the camera failed to capture a full image set due to the absence of the UBEC to regulate the peak power consumption of the camera. In the second survey, the power port in the DLS came loose, and therefore all images failed to have either the reflectance information or GNSS positions stored in the EXIF (exchangeable image file format) data. However, once the faulty DLS has been repaired by the camera manufacturer two successful surveys were completed at Penglaneinon on 26/06/18 and 02/08/18. At Ffridd Fawr, the terrain and difficult site access made surveying much more challenging. Given the terrain following parameter on Ardupilot was a relatively new function at the time of surveying, and that problems resulting from incorrect configuration of Ardupilot had already been witnessed the previous year, a cautious approach was taken to surveying the site. It was decided that surveying on the site should only take place once the UAV robustness was proven, and once the terrain following parameter had been fully tested at Penglaneinon where the slope was minimal. The

first survey attempt was undertaken in July 2018. However, the survey was unsuccessful as the UAV failed to launch correctly and suffered minor damage in the resultant landing. This was attributed to a freak headwind gust having occurred during the launch, which compounded with the pitch of the UAV being calibrated too low caused the UAV to be pushed down during launch. To remedy this, the UAV pitch was re-calibrated to a 5° incline, and the launch location was altered slightly to minimise the potential effect of sudden gusts. A second flight at Ffridd Fawr was conducted on the 29/09/18 and was successful. Before each successful survey, captures of the calibration reflectance panel were taken.

2.5.6 Image processing

There were 2450 captures taken for the Ffridd Fawr survey, resulting in a total of 12250 images being created (each capture featured five separate images - one for each multispectral band). For Penglaneinon, the survey taken on the 05/06/18 was chosen to be used for further analysis as it best coincided with animal tracking operations, and therefore was most representative of conditions at the time. This survey contained 3478 captures, resulting in 17,390 images being created. Image stitching was undertaken using Pix4Dmapper (Pix4D, Lausanne, Switzerland), which is based upon SFM algorithms. Processing steps included: (1) initial processing (keypoint extraction, keypoint matching, camera model optimisation, and GNSS geolocation); (2) point cloud construction (point densification); and (3) DSM, orthomosaic, reflectance map and index construction.

The main outputs required for this study were the digital terrain model (DTM), DSM, and 5 x reflectance maps (one for each spectral band - B, G, R, NIR, RE). The reflectance map represents a calibrated form of the orthomosaic, where the value of each pixel is adjusted to faithfully indicate the reflectance of the object. This calibration was achieved through using camera information (vignetting, dark current ISO etc.), sun angle and irradiance data from the DLS, as well as the calibration panel. The use of sun angle for reflectance calibration is cautioned against, however, as it can often lead to poorer results than when not utilised. It is recommended only when where considerable cloud variations (e.g. through intermittent cloud coverage) during a survey are present (Pix4D 2020). As a result, outputs with sun angle accounted for and not were compared to assess the best output. At both sites, the outputs were improved when sun angle was calibrated for.

At Ffridd Fawr, the total area covered by the final stitch was 155.66 ha, with a ground sampling distance (GSD) of 7.18 cm for the reflectance maps/DSM, and 35.9 cm for the DTM (5 x the DSM GSD). At Penglaneinon, the total area covered was 200.89 ha, and the GSD for the DSM and reflectance maps was 8.73 cm, whilst the GSD for the DTM was 43.65 cm (See appendices A.1 & A.2 stitching reports). All outputs were subsequently clipped to the extent of the sites using polygon shapefiles in QGIS (vers 3.4.15 ‘Madeira’). Minor areas of distortion along sections of the site

boundary at Penglaneinon were also masked out. The final study area remaining at Ffridd Fawr was 72.41 ha, and 109.46 ha at Penglaneinon. The coordinate reference system (CRS) for all outputs was WGS 84 /UTM zone 30N.

2.5.7 Image analysis

DTMs were used for the calculation for both slope and aspect and were also expressed directly as the elevation layers. Aspect and slope layers for both sites were created in QGIS using the raster analysis toolbox, with both algorithms being derived from the Geospatial Data Abstraction Library (GDAL) DEM utility. Both variables were expressed in degrees (°). NDVI was calculated at both sites using the standard formula:

Equation 1. Normalised difference vegetation index (NDVI)

$$\mathrm{NDVI} = \frac{(\mathrm{NIR} - \mathrm{Red})}{(\mathrm{NIR} + \mathrm{Red})}$$

Regarding the water availability layer, several steps had to be completed in order to produce a raster with corresponding values for each cell. Firstly, a series of vector polygons had to be created of the above ground water. It was decided that the permanence of these water bodies should be such that they would remain constant throughout the study period, and therefore temporary water bodies (e.g. puddles) were not considered. The methodology for deriving these polygons varied between the two sites, as a series of tools were utilised to achieve the best result for each site. At Ffridd Fawr, a normalised difference water index (NDWI) was created using the equation presented by (McFeeters 1996):

Equation 2. Normalised difference water index (NDWI)

$$\mathrm{NDWI} = \frac{(\mathrm{Green} - \mathrm{NIR})}{(\mathrm{Green} + \mathrm{NIR})}$$

The profile tool in QGIS was then used to derive a threshold from the NDWI for standing water (>-0.45), which was then incorporated into a new raster using the raster calculator. Additional water bodies not extracted by the NDWI threshold were identified visually using a combination of generated layers. These included a colour-infrared (CIR) layer (R = NIR band, G = red band, and B = green band) assembled using the virtual raster builder, and a canopy height model (DSM-DTM) visualised with a cumulative cut stretch featuring the full dataset extent. Identified water bodies were then polygonised and merged with the NDWI identified areas. At Penglaneinon, the NDWI threshold (>-0.45) created many false positives and was therefore only performed for the blanket bog area, using a clipped raster. The remaining water bodies on the site were extracted using the flow accumulation tool on the DTM in Arcmap (vers 10.5.0), and then merged with the blanket bog water

areas. Typically, distance from each cell to the nearest water body would then be calculated using 'Euclidian distance' - however, this does not account for topology. Given the elevation difference across the two study sites (especially Ffridd Fawr), distance was instead calculated using the 'path distance' tool in Arcmap, with the DTM inputted as the surface raster over which distance be calculated.

Standing shelter at Ffridd Fawr was considered to include trees, stone walls, derelict buildings (of which there were two), and a large rocky outcrop. These were extracted using a >0.18 m threshold derived from the canopy height model using the QGIS profile tool. Distance to each feature was then calculated using the path distance tool in Arcmap. Given the steepness of the slope in parts at Ffridd Fawr, topographic shelter was also considered a potential variable. This was calculated by creating a customised hillshade in QGIS. Hillshading is a 3D representation technique for visualizing terrain determined by a light source, and the aspect and slope of the elevation surface. The traditional method calculates the hillshade by creating an artificial illumination on the surface from one direction using specified altitude and azimuth properties to indicate the light source position. Given the westerly prevailing direction for the weather on site, the azimuth of the light was set to 270° (west) with the angle of the incoming light set to 45°. Under this specification, it was considered that any shaded areas would be indicative of varying levels of topographic shelter.

The final variable was the landcover classification. At Penglaneinon, this was completed within an accompanying MSc study (supervised on a day-to-day basis by the PhD candidate), using the Remote Sensing and Geographical Information System library (RSGISLib) (Bunting *et al.* 2014). Several steps were implemented including initial segmentation (using the Shepherd algorithm), training sample collection, initial classification, additional layer construction (e.g. slope, flow accumulation), and final classification (using the extra-trees classifier). A full explanation of the methodology may be found in Hinkley (2018). Accuracy assessment of the classification was completed using a random point assessment in Arcmap. The original UAV orthomosaic, and on-site ground-truthing data were used for verification. The overall accuracy of classification was 89% ($\kappa = 0.86$), with *Deschampsia cespitosa* deemed the most accurate (100%), and water the least (76%). The final output contains a combination of habitat based (peat bog), and species level (*Molinia caerulea*) classes. These classes were chosen as they were deemed most insightful regarding the management of site (i.e. management of the *Molinia*, and concern of over-grazing on the peat bog). *T. germanicum* (deergrass) was used a proxy for peatbog as it was observed to only cover the habitat area, and therefore the difficulties of separating class spectrally by habitat was not problematic in this case.

Unfortunately, due a combination of factors, the classification at Ffridd Fawr could not be completed. Firstly, unlike Penglaneinon, the vegetation at Ffridd Fawr was far more heterogeneous, with many

areas appearing to contain a number of micro habitats (e.g. bracken interspersed within semi-natural grassland) and therefore decisions on classes were more complex. This also meant that vegetation could not be easily classified according to spectral signature alone in the imagery and that thorough ground truthing would be required to verify any classification results. The COVID-19 outbreak and subsequent prohibition on fieldwork ultimately prevented this ground truthing from occurring. The methodology to be used was identical to that outlined above.

Once all the layers were generated, they were resampled to a 0.5 m GSD. This was due to there being a difference in resolutions between the layers derived from the DTM and reflectance maps, which meant the raster cells were not aligned (a key requirement for the further analysis on resource use). Resampling was completed in R (vers 3.3.6) using RStudio (vers 1.2.5). The 'raster' library was required (Hijmans *et al.* 2015). All outputs presented are the resampled layers (0.5m GSD).

2.6 Variables derived from UAV site imagery

Results presented include variables produced from the UAV derived imagery. All variables were successfully created and are demonstrated at a 0.5 m GSD resolution. Variables generated include elevation, aspect, slope, NDVI, distance to water, distance to standing shelter (at Ffridd Fawr only), topographic shelter (at Ffridd Fawr only), and landcover classification (at Penglaneinon only).

The elevation at Ffridd Fawr (Figure 2.16) is higher than Penglaneinon (Figure 2.17) at both the lowest and highest points, though much of the Penglaneinon site is situated between ~420 - 445m, whereas the elevation at Ffridd Fawr is more distributed evenly across the elevation levels.

Both sites are similar as regards their aspect (Figure 2.18 & Figure 2.19. North and east facing slopes predominate, however sections of south and west facing slopes may be found less frequently. Of the two most notable areas of the Ffridd Fawr site, the valley bottom section is south facing, whilst the main slope is north facing.

The slopes at both sites (Figure 2.20 & Figure 2.21) have been visualised along the same scale to better facilitate comparison. The hillside nature of the Ffridd Fawr site is evident from Figure 2.20**.** The north-eastern side of the slope contains the largest incline with ~29° pitch present for ~100 m of the slope. The contouring of the hill is also prominent, with a small gulley evident through the middle of the site. Contrastingly, most of Penglaneinon does not exceed a ~8° slope.

The semi improved grassland patches at the northern end of the Ffridd Fawr site are demonstrated clearly in Figure 2.22 to have the highest NDVI value and therefore to have the greenest vegetation for anywhere on the site. Likewise, at Penglaneinon (Figure 2.23), the sections of vegetation cut ~18 m previous to image acquisition clearly show a much greater NDVI then the surrounding areas. Though further investigation is required, the level of difference would indicate that legacy grazing has had an influence in maintaining the pasture's vegetative state.

The maximum distance from water at both sites is similar, though slightly greater at Penglaneinon (+67 m further then at Ffridd Fawr). As shown in Figure 2.25, the water bodies at Penglaneinon are placed fairly centrally within the site. At Ffridd Fawr (Figure 2.24), the site is intersected by a stream running through the middle of the site in the gully (Figure 2.20), however the remaining water bodies predominate in the northern end of the site, with little in the south. All water bodies identified at both sites were either small tributaries/streams. The largest water body at either site was a narrow river (~2 m wide) which runs along the valley bottom at the northern end of the Ffridd Fawr site.

Standing shelter at Ffridd Fawr (Figure 2.26) consisted of a stone wall, two derelict buildings (one with partial roof intact), a small number of trees, and a rocky outcrop at the southern end of the site. As is evident from Figure 2.26, all standing shelter is located near the eastern side of the site, with

none on the western side. This is similar for topographic shelter. As demonstrated in Figure 2.27, the majority of the steeper slopes which represented the most sheltered areas, were also present in the eastern areas of site, predominately between the northern and southern standing shelter. As expected, there is little topographic shelter towards the top of the site, approaching the western side.

The vegetation classification at Penglaneinon (Figure 2.28) highlighted the extent of the *Molinia* coverage on the site. A total of 76.5% of the area was covered in *Molinia*, whilst 18.9% of the area was composed of a mixed grass sward. The peatbog covered 3.4% of the site, whilst water and *Deschampsia cespitosa* accounted for 3.4% and 0.1% respectively. It is notable that all of the previously cut areas are classified as mixed grass species, which may indicate that cutting (and potentially legacy grazing) have altered the species composition of those sections. However further ground truthing would be required to validate this. Unfortunately, due to the COVID-19 outbreak, this could not occur within the current study.

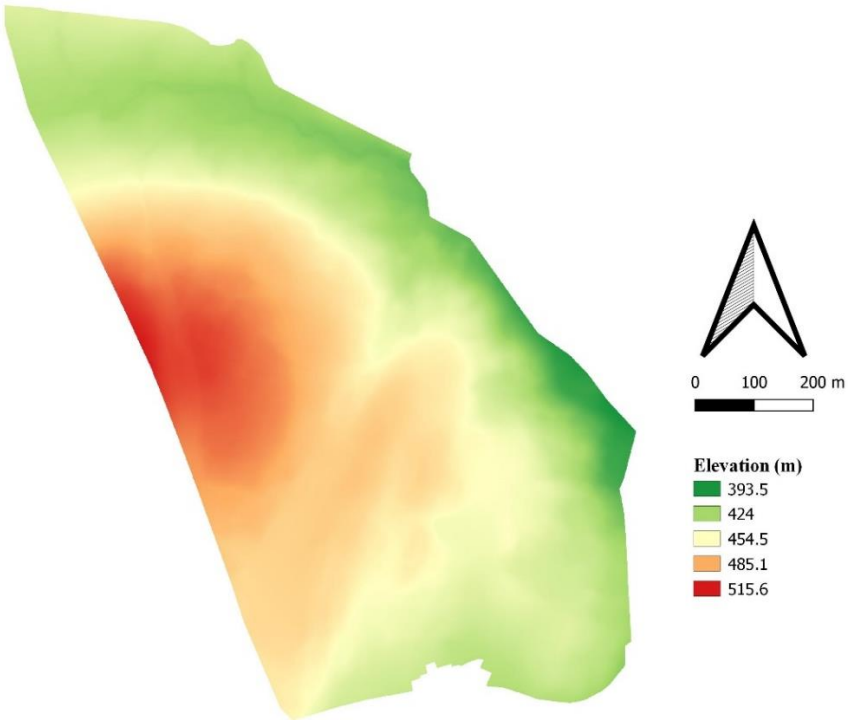


Figure 2.16 Elevation layer for Ffridd Fawr represented by the site digital terrain model (DTM). A local histogram stretch was used to present the layer. Colours represent specific levels on a continuum.

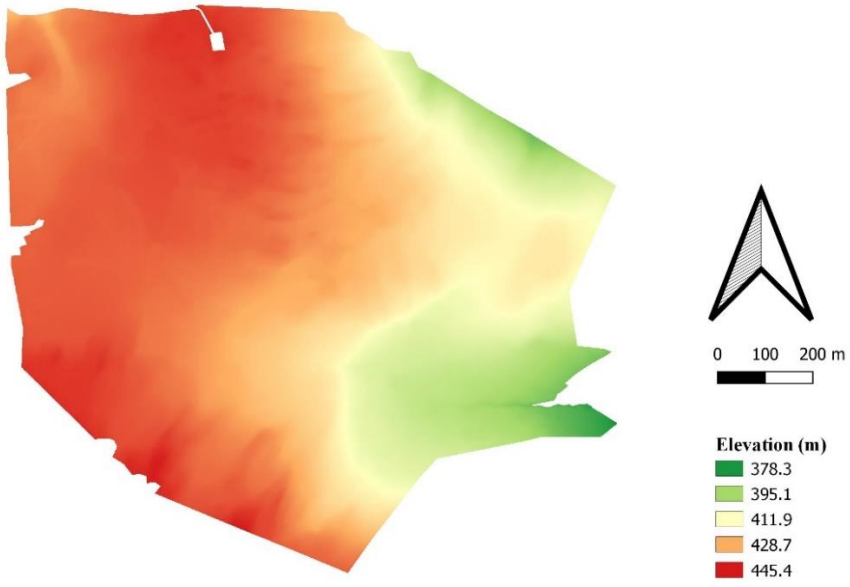


Figure 2.17 Elevation layer for Penglaneinon represented by the site digital terrain model (DTM). A local histogram stretch was used to present the layer. Colours represent specific levels on a continuum.

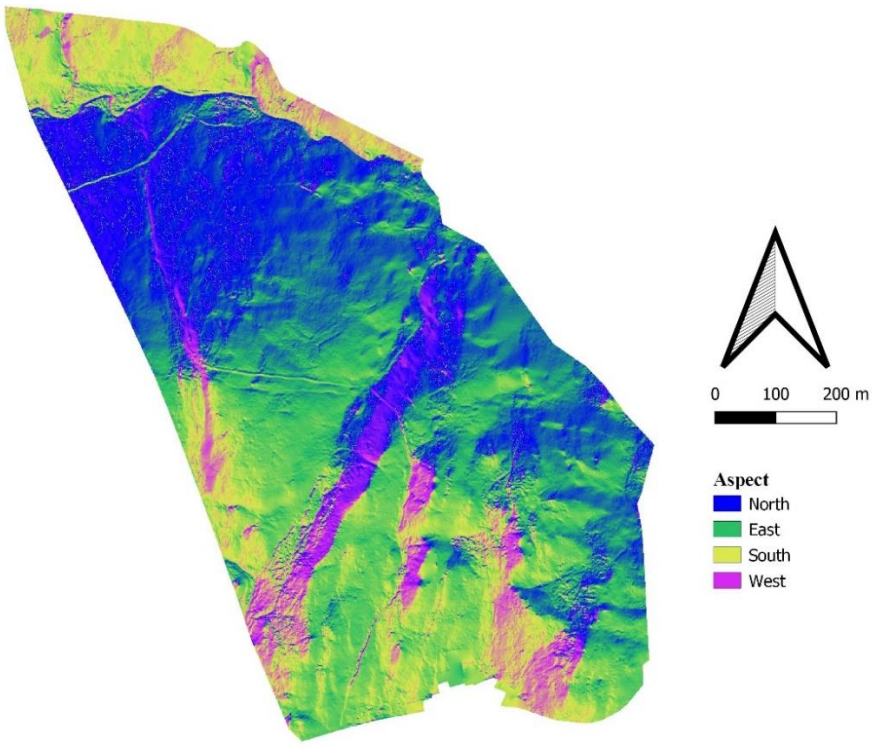


Figure 2.18 Aspect layer for Ffridd Fawr constructed from the site digital terrain model (DTM). Values for the layer represent the Azimuth (°) of each cell. The four cardinal directions (N, E, S, W) were symbolised using a custom colour ramp in QGIS (North = 0° & 360°, East = 90°, South = 180°, West = 270°). A local histogram stretch was used to present the layer. Colours represent specific levels on a continuum.

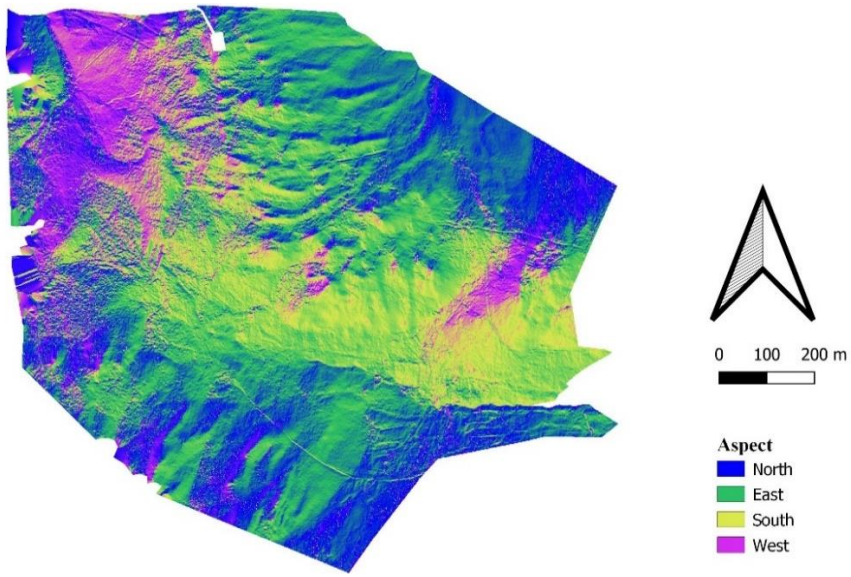


Figure 2.19 Aspect layer for Penglaneinon constructed from the site digital terrain model (DTM). Values for the layer represent the Azimuth (°) of each cell. The four cardinal directions (N, E, S, W) were symbolised using a custom colour ramp in QGIS (North = 0° & 360°, East = 90°, South = 180°, West = 270°). A local histogram stretch was used to present the layer. Colours represent specific levels on a continuum.

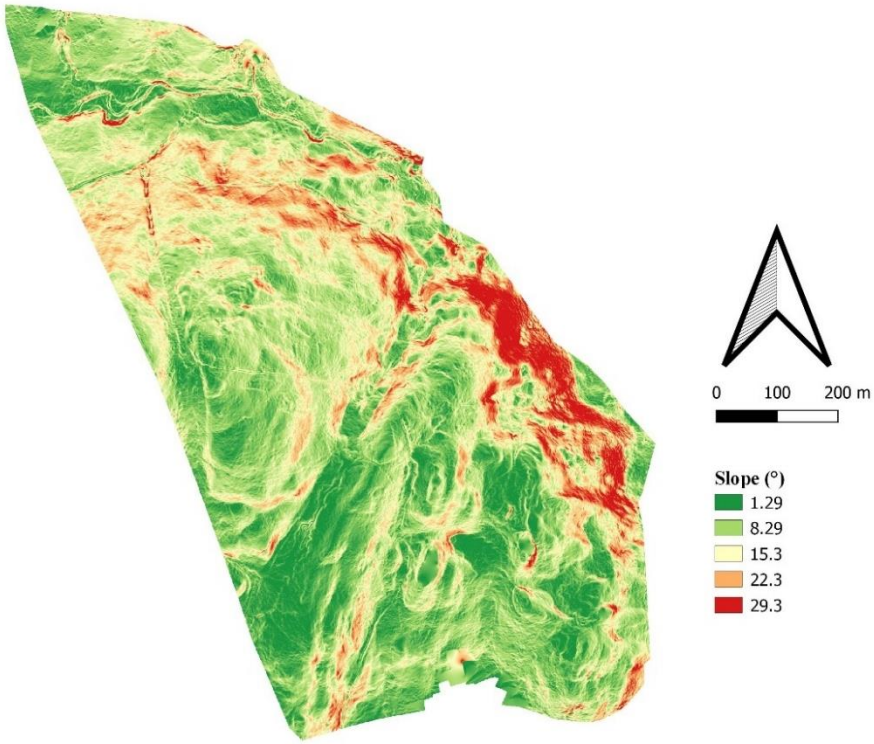


Figure 2.20 Slope layer for Ffridd Fawr. A local cumulative cut stretch was to present the layer, as it better exemplified the changes in slope across the site. The true minimum and maximum slope for the site was 0.01° and 49.3°. Colours represent specific levels on a continuum.

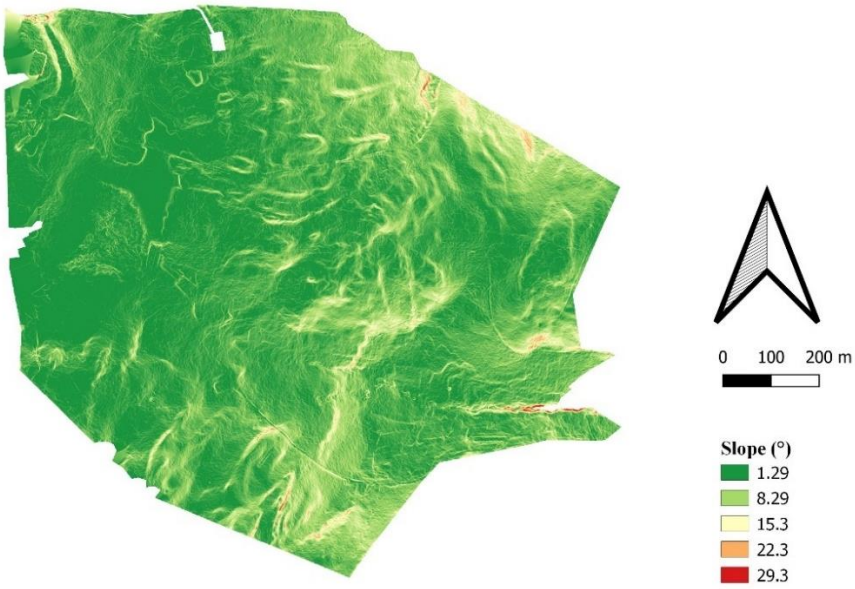


Figure 2.21 Slope layer for Penglaneinon. The values used for the stretch were matched with that of Ffridd Fawr to allow true comparison of the slopes between sites, thereby avoiding an exaggeration of the slope at Penglaneinon. The true minimum and maximum slope for the site was 0° and 36.9°. Colours represent specific levels on a continuum.

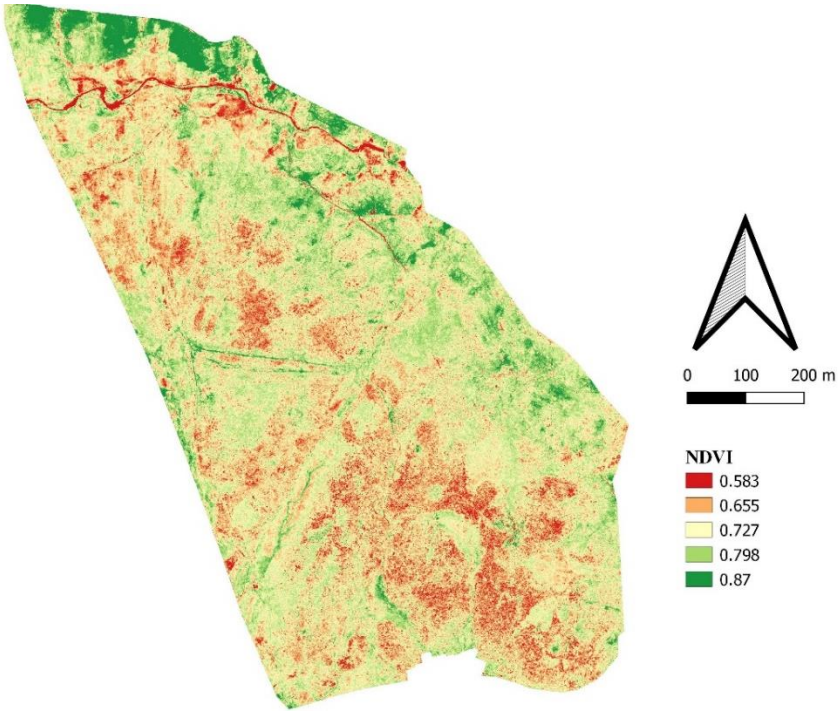


Figure 2.22 Normalised difference vegetation index (NDVI) layer for Ffridd Fawr. A local cumulative cut stretch was to present the layer, as it better exemplified the changes in NDVI across the site. The true minimum and maximum NDVI for the site were - 0.090 and 0.972. Colours represent specific levels on a continuum.

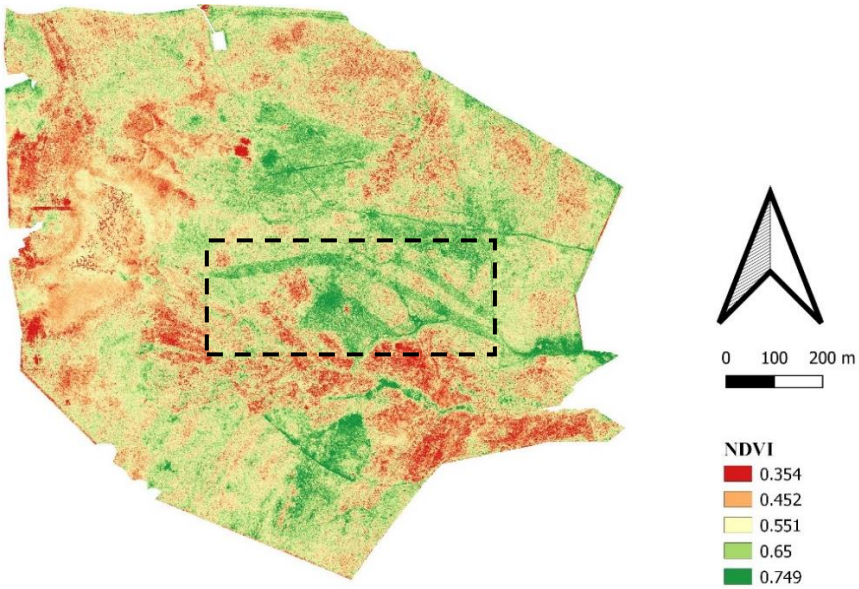


Figure 2.23 Normalised difference vegetation index (NDVI) layer for Penglaneinon. A local cumulative cut stretch was to present the layer, as it better exemplified the changes in NDVI across the site. The true minimum and maximum NDVI for the site were 0.293 and 0.896. Black box denotes vegetation patches cut 18 months previously. Colours represent specific levels on a continuum.

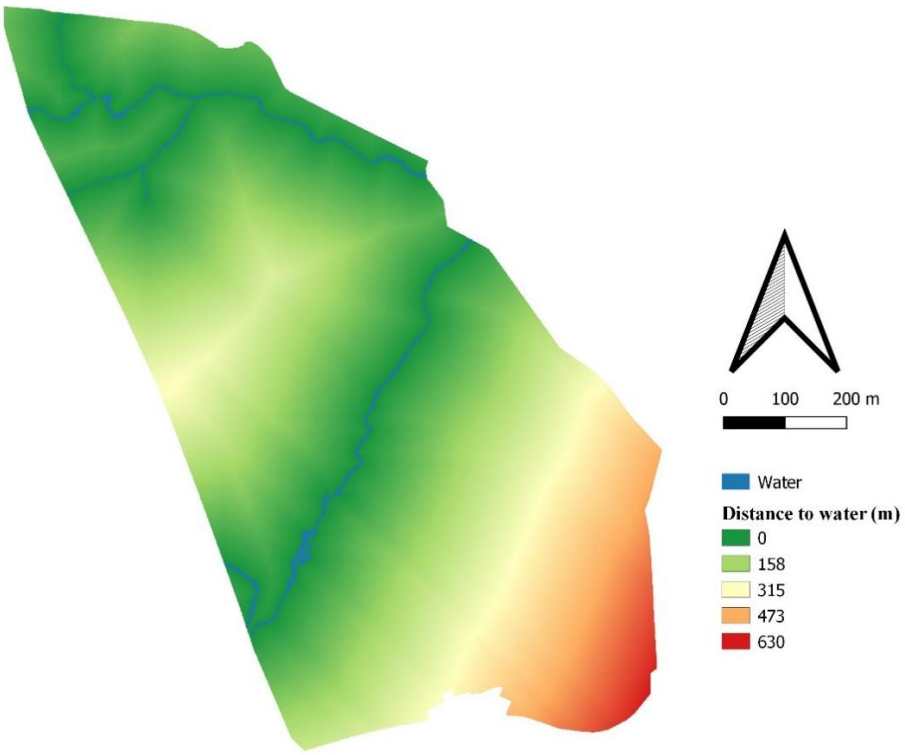


Figure 2.24 Distance to water layer at Ffridd Fawr site, with identified water bodies overlaid. Distance across site terrain was calculated using the site digital terrain model (DTM). A local histogram stretch was used to present layer. Colours represent specific levels on a continuum.

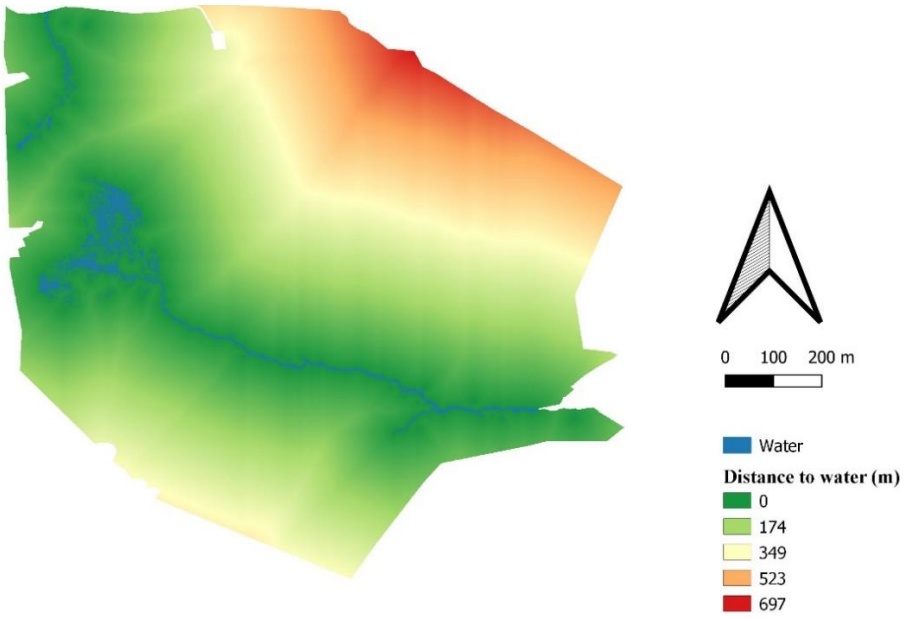


Figure 2.25 Distance to water layer at Penglaneinon site, with identified water bodies overlaid. Distance across site terrain was calculated using the site digital terrain model (DTM). A local histogram stretch was used to present layer. Colours represent specific levels on a continuum.

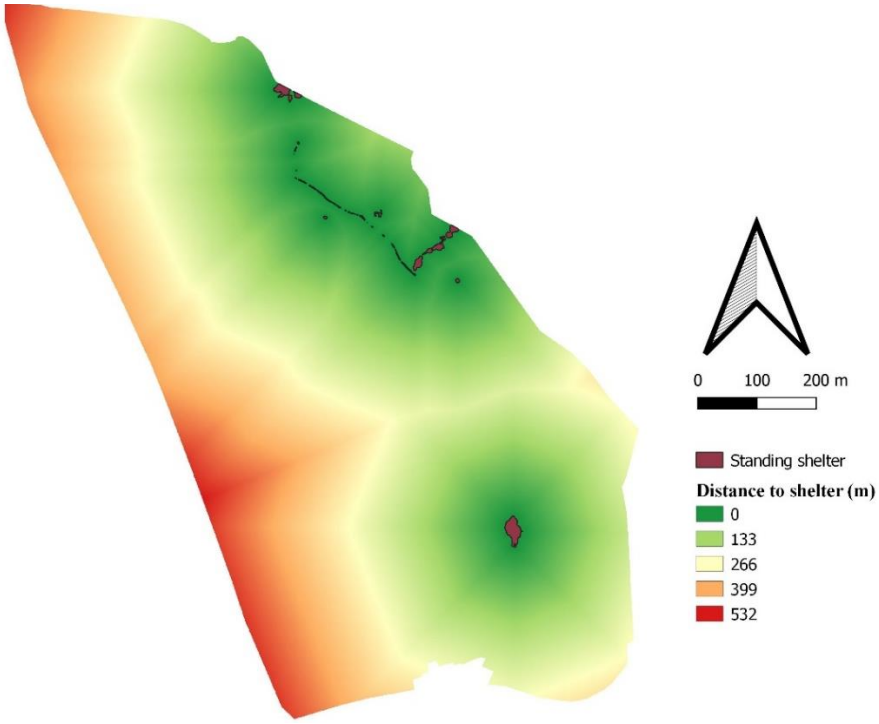


Figure 2.26 Distance to standing shelter layer at Ffridd Fawr site, with identified shelter overlaid (trees, stone walls, derelict buildings, and a rocky outcrop. Distance across site terrain was calculated using the site digital terrain model (DTM). A local histogram stretch was used to present layer. Colours represent specific levels on a continuum.

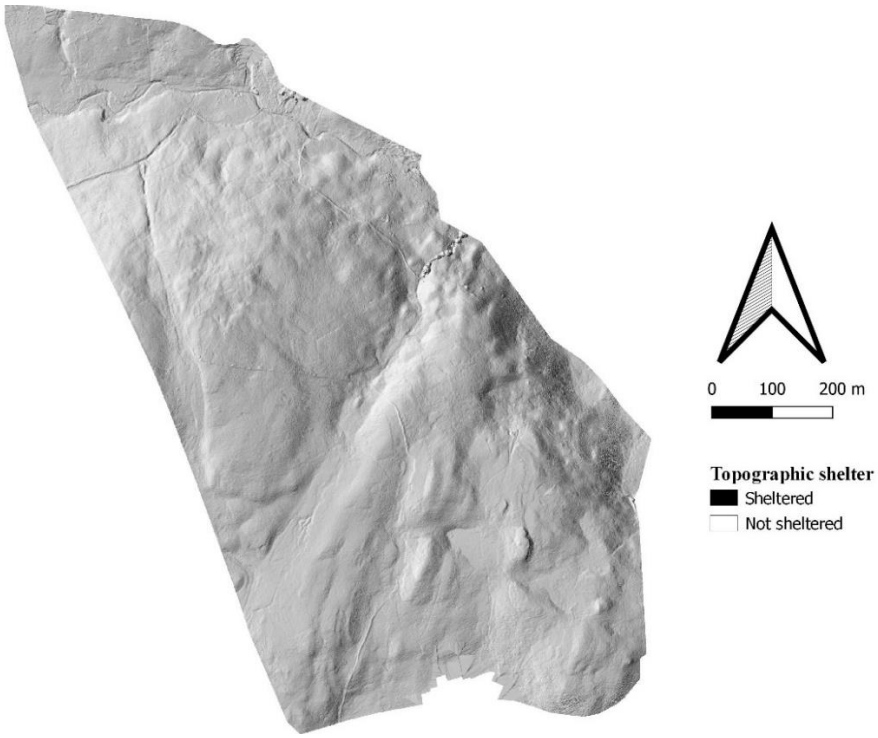


Figure 2.27 Topographic shelter at Ffridd Fawr represented by a customised hillshade. The azimuth of the light was set to 270°(west) to match the prevailing wind on site, and with the angle of the incoming light set to 45°. A local histogram stretch was used to present layer. Colours represent specific levels on a continuum.

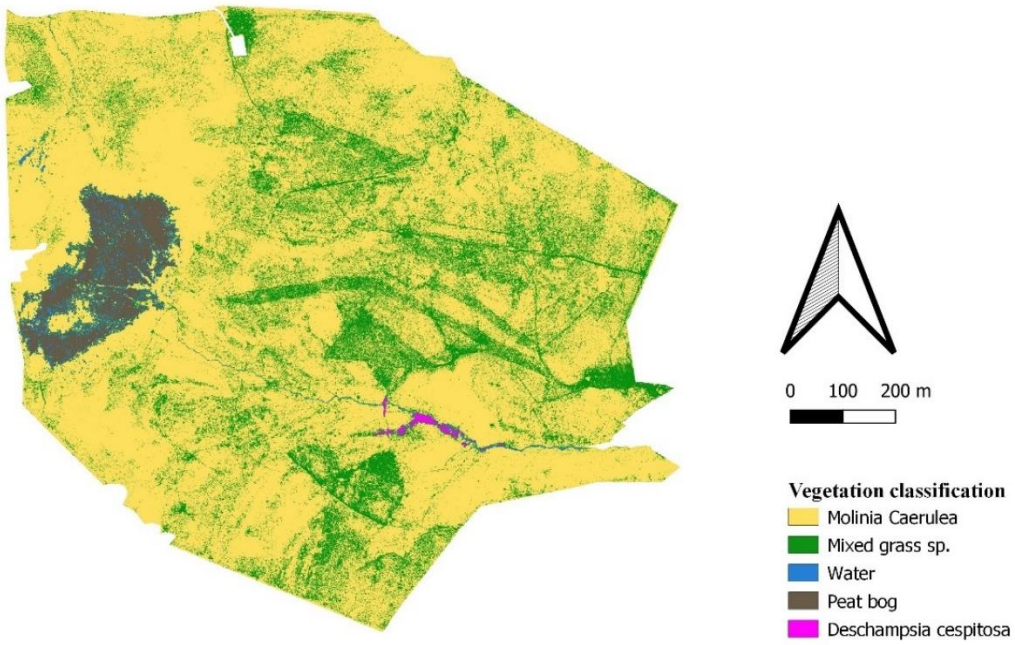


Figure 2.28 Landcover classification for Penglaneinon. The overall accuracy of classification was 89% ($\kappa = 0.86$). Colour represent categorical levels.

2.7 Chapter outcomes summary

In summarising the results, it is clear that whilst a number of similarities exist between the sites, the level of differences highlight why each particular site should first be characterised in detail. In this case, the results obtained from the profiling are sufficient to enable the two study sites to be compared and contrasted with confidence.

The successful generation of the resources and factor layers likely to affect sheep on both sites ultimately validates the use of UAV derived imagery for this purpose. The high resolution of outputs produced provide a considerable level of detail not currently accessible by any other data collection means. The use of UAVs for peatland mapping is still in its infancy, with all identified studies in the literature being conducted over smaller study areas then the sites in this study (Lovitt, Rahman & McDermid 2017; Rahman *et al.* 2017; Palace *et al.* 2018; Beyer *et al.* 2019) (the largest survey area was 61 ha in Lovitt, Rahman, & McDermid (2017)). Despite the study areas being larger in the current study, the quality of outputs produced, and level of accuracy attained (89%) compare favourably with published results. Beyer *et al* (2019) achieved 87% vegetation classification accuracy, Palace *et al (2018)* 68%, and Knoth *et al* (2013) 91%. Furthermore, it has been demonstrated in this study that such mapping can be completed at a low cost using a self-build UAV design. The approach and subsequent methodology utilised in this study therefore presents a compelling option when compared to existing approaches.

However, constructing a self-build UAV requires a competent level of experience in robotics and aviation theory, and therefore may not be suitable for many inexperienced users unless guidance can be sought. A sufficiently long testing period (varying according to expertise, but likely at least ~2-3 months) should also be accommodated, so as not to risk UAV damage and failures in collecting critical data when surveying commences. The use of such a UAV design within a project with minimal preparation time is therefore not recommended. That said, when time and experience are available, the level of customisation available for self-build designs means the UAVs can be tailored to the specific needs of a project. The promotes a high level of efficiency within the data collection process, by maximising resource use and providing a high level of confidence in the data collected. Finally, the use of commercial UAVs in many cases does not negate the chances of mechanical issues, and an advantage of using a self-build design is that users will more likely be able to recognise any faults sooner and address them more confidently. This can be vital during surveys where the amount of time available to complete a mission is critical (e.g. because of weather), and therefore being confident in the mechanics of system utilised can make the difference between successful data collection and failure.

Though the use of UAVs within the context of this study was prudent, it must be reiterated that other studies assessing larger areas or requiring a lower GSD resolution could well be better suited using other options (such as satellite imagery) (Frick *et al.* 2011; Bourgeau-Chavez *et al.* 2017). Furthermore, it may well be that whilst UAVs be used for an initial site survey, follow up data could be collected via other means if only a few variables needed measuring. For example, in the case of the sites in this study, factors such as slope and aspect are fixed, whilst others such as shelter and landcover classes are also unlikely to change in a short amount of time. Therefore, a user could use data such as Sentinal-2 derived imagery to monitor a changeable factor such as NDVI change over a season (Raynolds *et al.* 2015), or manually assess a site to digitise how water bodies varied. However, in another scenario, a user may well wish to attempt monitoring vegetation height, which may be better undertaken by a UAV (Forsmoo *et al.* 2018).

In summary, the choice of data collection tool for characterising a site should be determined by the precise variables wished to be measured, with the reasons why clearly stated (i.e. for resource utilisation analysis in the case of this overall study). These variables in turn should only be chosen once an exhaustive profile of a target site has been undertaken, with any pre-existing data considered first. Once this has been undertaken, then the appropriate sampling method(s) may be selected, and in many cases, a combination of sampling approaches may be appropriate. In short, allowing the specific biological questions to dictate the methodology rather than vice versa.

CHAPTER 3

DEVELOPING LOW COST APPROACHES TO TRACKING LARGE NUMBERS OF ANIMALS

3.1 Introduction

3.1.1 The importance of movement ecology

As a scientific discipline "movement ecology" has been crucial in progressing our knowledge on many ecological and evolutionary processes. An understanding of animal movement can inform a wide range of topics from animal physiology, disease spread, population dynamics and community ecology, to nutrient cycling, habitat selection, wildlife management and conservation (Calabrese, Fleming & Gurarie 2016). Within the context of the current study, its importance can best be surmised by Perotto-Baldivieso *et al* (2012) who state that *'categorizing animal movement is key to quantifying an accurate relationship between the spatial distribution of organisms and the resources they require for survival'*. Having an accurate reflection of this relationship is therefore key to understanding the animal-environmental interaction on a given site, and only through this can manipulation towards a more desired utilisation of the surrounding environment be achieved.

3.1.2 The biological objectives of the study

Interpreting animal movement correctly is seldom straightforward and is often considered a key challenge in conservation biology and ecology. Movement after all is a complicated process that is influenced by a number of different biotic and abiotic factors, whether it be the internal state of an animal, its physiological constraints, or the environment it operates within (Fleming *et al.* 2014). It is therefore important to consider the specific biological objectives of a given study carefully (Perotto-Baldivieso *et al.* 2012), starting with the optimum analysis options for answering the designated questions, to the experimental design of the tracking study and the tracking technologies most suited to delivering the required data.

The end goal of the overall study is to be able to rank the environmental variables affecting animal distribution on each site. Therefore, the requirement for this is that any tracking is representative of all the animals on each site. In Chapter 1, the second aim outlined was to '*track animals on each site, and test their respective supposed uniformity of grazing intensity*'. Analysis of the tracked animals should therefore include reference to the use of LSUs as well as other estimations of spatial use (e.g. home range analyses). Finally, the central research question for PhD study *('Can recent developments in sensing technologies and/or processing capability be harnessed to improve the evidence-base for conservation decision-making'*) prompts an investigation as to the tracking technology available and their suitability for the tracking required.

3.1.3 The need for accessible analyses

With key objectives identified (Table 1.2), it is possible to focus on the analysis options which would be most prudent in delivering the required outcomes. However, underpinning these objectives is the desire to identify analysis options readily accessible to non-academics. After all, many habitat conservation activities are undertaken by NGOs and charities where management decisions are carried out by land managers, farmers, park wardens and rangers. Complex analyses which require significant computer programming experience therefore may be restricted in their use, misinterpreted in methods requiring interpretation of the results (Torres *et al.* 2017), or executed with unknown or false assumptions employed in the method(s).

Identifying solutions for these individuals which are easy and quick to perform, remove complex interpretations from the outputs, and clearly explain the assumptions by which they are executed, have the ability to provide succinct and accurate information for evidence-based land management decisions. That said, this project did not intend to diminish the role of rigorous statistical methods. To fully understand animal-environment studies using statistical evaluation, proper analyses are required to optimize and accurately understand ecosystem processes (Perotto-Baldivieso *et al*. 2012). Therefore, with regards to the overall aim within this study *('rank the environmental factors affecting the spatial distribution of animals on given site'*), a rigorous statistical approach is necessary to correctly interpret the output (which will be addressed in Chapter 4). However, in this chapter there remains an opportunity to explore simple, quick analytical options which could be used at a practical level that could be useful in informing management decisions.

One of the more frequently employed methods used in movement ecology are home-range analysis techniques. Since Burt's (1943) early conceptual definition (*"...the area traversed by the individual in its normal activities of food gathering, mating, and caring for young. Occasional sallies outside the area, perhaps exploratory in nature, should not be considered..."*-) the home range concept has gone on to appear throughout varying aspects of the ecology and evolution literature (Péron 2019). The ease with which many of the techniques can be implemented (Fieberg 2007), and the simple but effective ways they can be visualised make it appealing in a number of scenarios. Within this study, it presents as an effective way of measuring grazing intensity across each site, for a collective group of animals (e.g. to allow comparison with LSU estimations) and on an individual basis (e.g. to view variations in hefting). It can also be used to identify landscape features important for behavioural activities such as finding water, food, and shelter (Kie *et al.* 2010; Cohen *et al.* 2018) However, its popularity and ease of use has created some discord with regard to its implementation, with concerns raised that is often misused and regularly prone to either over/underestimation of area. Indeed, even the use of 'home range' as a term has been debated (Johnson *et al.* 2008), with Péron (2019) suggesting that the phrase be restricted to studies that are performed at demographically relevant time

frames (typically multiple seasons/years), and instead use ‘movement amplitude’ for shorter durations (e.g. a single season). However, the concern most often expressed with home range estimators is that of auto-correlated data.

3.1.4 The concerns and considerations in GNSS (Global Navigation Satellite System) tracking studies

3.1.4.1 Autocorrelated data

Autocorrelation in simple terms can be defined as the degree of similarity between observations recorded at different time points (Odland 1988). The main reason it is problematic within tracking studies is because sequential observations are not independent in time or space, which violates assumptions for statistical interference (Boyce *et al.* 2010).

Before GPS (Global Positioning System) tracking studies were predominately undertaken using radio telemetry technology, and therefore sequential points could often satisfy the need for being independent. But with the advent of GPS, the capability to undertake automated data collection has often generated large amounts of position and activity information at a high frequency (Ungar *et al.* 2005). While this ability to collect detailed information on animal movements has provided numerous opportunities for analysing animal behaviour, the volume and frequency of data collected has caused the issue of autocorrelation to become far more prevalent (Perotto-Baldivieso *et al.* 2012). Though autocorrelation is a serious concern for tracking studies, it can often be managed effectively either through data thinning procedures, or more recently through certain statistical analyses which do not assume spatial/temporal independence between points. An example of this for home range estimation is auto-correlated kernel density estimation (AKDE) (Péron 2019).

3.1.4.2 The cost of GNSS tracking and other implications on experimental design

The design of GNSS tracking studies often raises more concern then that of autocorrelated data, chiefly due to the low numbers of animals typically tracked (Allan et al. 2013; McGranahan, Geaumont & Spiess 2018), and therefore the reliance on the positional information of only a few individuals being used to represent whole herd/group movements (Anderson, Estell & Cibils 2013). This is primarily attributed to the high individual cost of GNSS commercial loggers (Perotto-Baldivieso *et al.* 2012) with prices at the time of writing being anywhere from £500-£2000 per unit depending on specification. In the current study, this cost was particularly problematic. Given the number of animals contained on each site (~200), simply collaring a few animals would almost certainly underestimate the respective site utilisation distributions (UD) – the probability density that an animal is found at a given point in space. Addressing this problem was therefore imperative in order to reliably satisfy the overall objectives of the project. A second technical concern was that the animals on both sites are only handled twice in the sampling period. Given this time frame stretches

from mid-May to mid-September, it is feasible that any deployed tracking equipment will have to last ~50 days on a single battery charge. As GNSS loggers typically reside in a sleep state between recordings in order to preserve power, this means that any recording schedule implemented would either have to be conservative in nature for the retention of battery life between charges, or take a higher intensity approach which results in a shortened tracking period. However, the latter approach would obviously cause gaps within the data between tracking sessions. This is not an isolated problem within tracking studies, with Cohen *et al* (2018) cautioning that researchers must optimise battery life relative to the data needed to address hypothesis of interest.

3.1.5 Chapter objectives

The primary objectives of the current study necessitated a comprehensive tracking strategy for outputs of the analysis to be robust and meaningful. Specifically, sufficient individuals had to be tracked for grazing intensity and ranking of resource use on each site to be truly representative. The first part of this chapter will focus on the development of a new low-cost radio tracking solution using unmanned aerial vehicles. The second part will then assess how open source/low cost GNSS loggers can be utilised to maximise their output potential with regard to their battery management, plus ability to offer mixed frequency data for different analyses options. Given that the intention is to make this as accessible as possible to practical users, simple analyses options (including new software developed for behavioural inference) will be demonstrated that can be easily implemented. Finally, ways in which the outputs of these analyses, both individually and combined, can help inform on the supposed uniformity of grazing intensity will be discussed.

3.2 A bespoke low-cost system for radio tracking animals using multi-rotor and fixed-wing unmanned aerial vehicles (UAVs) (Published, '*Methods in Ecology and Evolution*'-see appendix A.4)

3.2.1 Introduction

3.2.1.1 The motivation and reasoning behind this work

The hardware development in this section primarily addressed a particular limitation of equipment that was currently available. Because of the high individual cost of commercial GNSS loggers, this typically leads to a low number of collared animals (usually less than 10 collars, and often as low as 1 or 2) being used to represent whole herd distributions (Anderson, Estell & Cibils 2013). Given the sheer number of animals involved in our study, and the local knowledge that hefting wasn't thought to be uniform on the sites, it was deemed unwise to solely rely on GNSS loggers to represent the true distribution on site.

Initially, other tracking technologies aside from GNSS were explored, with both commercial and open-source solutions being considered. The most obvious and readily used alternative was radio frequency (RF) tracking. Whilst radio RF tracking tags cannot provide the continuous tracking capability of GNSS equipped trackers, they are inexpensive and can be extremely small and lightweight, thereby allowing large number of animals to be tracked albeit at lower spatial precision and frequency. Furthermore, advances in the autonomous capability and payload capacity of unmanned aerial vehicles (UAVs) has led to them being increasingly utilised and explored as potential data collection platforms in ecological surveying and monitoring (Hodgson *et al.* 2018). The ability of UAVs to travel long ranges quickly (particularly in the case of fixed-wing UAVs) whilst offering greater likelihood in the line-of-sight of target animals (Körner *et al.* 2010) provides advantages over conventional methods of radio tracking on the ground. It was reasoned that increased accuracy could well be one of these advantages. Finally, it could be argued that the use of UAVs would likely lead to low temporal resolution data, given that opportunities for UAV flights are fundamentally dependent on local meteorological conditions. However, this does provide an advantage in the fact that the data gained would almost certainly be spatially and temporally independent (i.e. not autocorrelated). Considering this, but also the fact whole herds could feasibly be tracked, it was proposed that this data be ideal for monitoring overall grazing intensity and cumulative site use over a tracking period. Given the objective to measure the supposed uniformity of grazing for each site, incorporating this data into analyses such as home range estimators, or other tailored site-use intensity measurements, would therefore likely present a robust solution to this objective.

3.2.1.2 An introduction to UAV based radio tracking systems.

Over the last few years, researchers have begun to explore the potential benefits of UAV-based radio tracking systems (hereon referred to as UAVRTS) compared to conventional methods. However, as Shafer *et al* (2019) note, many of the presented systems exist primarily as proof-of-principle concepts. The prime focus in most of these studies is the refinement of the localisation methods employed. Whilst this may be valuable in considering potential hardware configuration options, there remain sizeable knowledge gaps within the subject area that have delayed the development of field-ready systems. Firstly, there has been very limited testing on animals, with tagging to date almost exclusively restricted to avian species only (Cliff *et al.* 2015; Tremblay *et al.* 2017). Furthermore, many studies have been limited to single tag testing (Körner *et al.* 2010; Dos Santos *et al.* 2014; Bayram, Stefas & Isler 2018; Shafer *et al.* 2019), and thus their ability to track movements when multiple animals are tagged remains unknown. Furthermore, none of the studies have utilised or tested their systems on fixed-wing UAVs. Given that fixed-wing UAVs offer vastly superior range, flight speed and endurance compared to multi rotor platforms, there is an opportunity to greatly expand the capability of UAVRTS by using such a platform.

3.2.1.3 How this study differed

The novel system reported in this study features a fully custom-made active radio-frequency identification (RFID) tag and receiver system suitable for both fixed-wing and multi-rotor UAVs. The electronic components chosen are purposely low cost with the goal of making tagging greater numbers of animals more affordable. Unlike previous studies where existing commercial tags have been used or modified, a bespoke tag specifically designed for detection by a UAVRTS is presented. Whereas most existing tags continually transmit when activated, these RFID tags remain in a dormant state, with a brief listening period occurring every 6 s. Tag responses are only elicited when a tag exciter trigger located on the UAV comes into operation, thereby saving considerable battery life. The receiver system is also contained within a single printed circuit board (as opposed to the multi-component set-ups utilised within previous studies); which substantially reduces the overall weight as well as the likely mean time between failure (MTBF).

Previous UAVRTS have focused on incorporating and modifying either direction (e.g. direction of arrival) or range-based techniques (e.g. received signal strength) as methods of locating tags. This study explored an alternative localisation method. Using grid flight mission functionality available in both open-source and commercial autopilot systems, estimated tag locations were derived by a simple mean coordinates calculation (Figure 3.1). The assumption of equal coverage of the surveyed area (provided by the flight grid), as well as the notion that the grid exceeds the range of the tag (i.e. so

estimated locations are not simply the centre of the grid) are central to accurate tag location by this method. By flexibly altering the transmission power of the tag trigger exciter depending on the size of the grid employed, signal loss at grid edge regardless of the situation could be ensured.

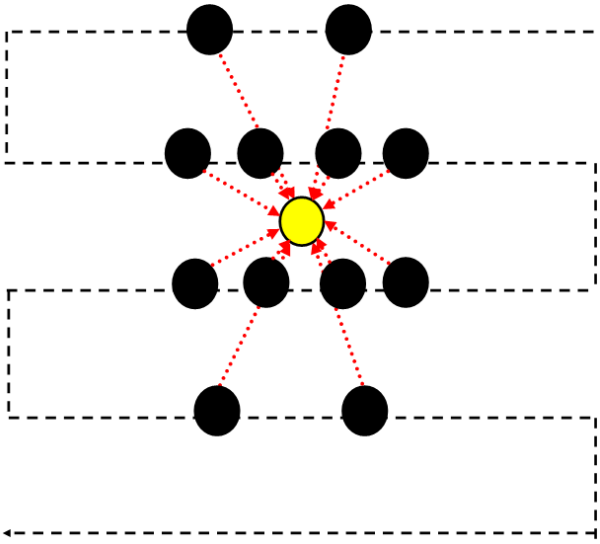


Figure 3.1 Locating tags using mean coordinates. Black line = UAV flight grid. Black dots = receiver location when tag transmission was detected. Yellow dot = estimated tag location based on mean coordinates of black dots.

3.2.1.4 The objectives regarding our UAVRTS

At the time of writing, this method of localisation is undocumented in the literature for UAVRTS. It was therefore deemed prudent to explore the level of accuracy deliverable, and the operational considerations that could affect it. Flight speed was considered to be the key parameter of interest, with specific exploration focused on: 1) the effect it had on the on the number of hits (tag responses) received; 2) how this ultimately affected the accuracy of a determined tag location; 3) assessing the cross-compatibility of the UAVRTS to function on both multi-rotor and fixed-wing UAVs; and 4) measuring the performance and reliability of the system with tags placed on animals, so the real-world applicability of this system would be tested.

3.2.2 Methods

3.2.2.1 RFID tag system design

The main components in each RFID tag (Figure 3.2) were a PIC10F206 microcontroller (Microchip Technology Inc, Chandler, Arizona, USA) and a HopeRF RFM69W radio transceiver (Hope Microelectronics co., Ltd, Nanshan District, Shenzhen, China) operating in the 868 MHz band. These were mounted on a custom-printed circuit board with integrated antenna and were powered by a single cell 60 mAh lithium polymer (LiPo) cell. The microcontroller was programmed to wake the radio module approximately once every 6 s for a period of approximately 2 ms. During this 2 ms period the radio module detected the signal from the trigger module (if one was present), switched to a different radio channel and responded using a simple medium access control delay mechanism with a radio packet containing a unique identifier for the tag. The response was transmitted three times, again using a simple medium access control delay mechanism to help reduce collisions between packets from different tags. In the absence of a signal, the microcontroller returned all components to a low power sleep state until the next listening window 6 s later. Predicted battery life in the absence of a trigger signal is over one year, but each transmission triggered will reduce the battery life by around one hour. Responses from the tags were recorded by the UAV mounted receiver module which used a HopeRF RFM69W radio transceiver, a Quectel L86 GPS receiver (Quectel, Xuhui District, Shanghai, China) and an ATMega328P microcontroller (Microchip Technology Inc, Chandler, Arizona, USA). The microcontroller decoded the packets received from the RFID tag and saved the tag unique identifier, latitude and longitude of the receiver and timestamp to a removable microSD memory card. With a clear line-of-sight and using a 10mW transmitter power in both directions (from trigger to tags and tags to receiver), the range achievable varied between 500 m and 800 m. Total weight for the system on the UAV was 195 g; this included; receiver box (115 g), trigger (80 g), batteries, and cable ties, etc. Each RFID tag weighed 9 g. At the time of writing, the estimated cost was £160 (£135 for receiver, £25 for trigger) with each RFID tag priced at just under £12.

Figure 3.2 Custom built radio-frequency identification tag.

3.2.2.2 UAV set up and RFID tag system integration

Multi-rotor: The platform was a DJI Phantom 3 professional (DJI, Nanshan District, Shenzhen, China). The receiver and tag trigger were mounted onto opposing ends of a 1 m long fiberglass rod, which in turn was cabled tied to the two landing stands on the drone (Figure 3.3). The UAV was operated autonomously using the PIX4D capture app (PIX4D, Lausanne, Switzerland) on an Iphone 5S (Apple Inc., Cupertino, California, United States).

Fixed-wing: The UAV set-up was similar to that of (Ryan *et al.* 2015). The UAV airframe was a Skywalker X8 (Skywalker; Hubei, P.R. China). Autonomous flight capability was used (http://ardupilot.com/); this provided flight stabilisation, altitude control (including terrain following), and GNSS navigation. The tag trigger and receiver were located on opposite wing tips, with each accompanied by a single rechargeable 300 mAh LiPo cell as a power source, which could provide ~2 hours of use (Figure 3.3). The receiver was encased in a small plastic container wrapped with aluminium foil, except for directly above the GNSS module, as initial testing revealed considerable radio interference from the fixed wing UAV avionics. Additional shielding was also fitted over the speed controller and electrical cables to the motor. The receiver case was bolted onto the wing tip, whilst the tag trigger was attached using cables ties, and the join further strengthened using cross-weave tape.

Figure 3.3 DJI Phantom 3 Pro with radio-frequency (RF) system mounted along a plastic rod attached to under carriage (top left). Skywalker X8 with RF system attached on wing tips (top right). RF receiver mounted inside foil wrapped box on X8 wing tip (bottom left). Tag trigger mounted on X8 wing tip (bottom right).

3.2.2.3 Static accuracy and the effect of UAV flight speed

Multi-rotor: Eighteen RFID tags were split into three groups which were each placed at three different locations ~200 m apart (n = 6 per group), with every tag within each group equally spaced within a 1 m^2 area. Additionally, two GNSS loggers (Ystumtec Ltd, Aberystwyth, United Kingdom) were present in each group to provide a reference location for each group. The flight grid consisted of a four-line grid encompassing a 650 x 230 m area. Twelve flights were conducted in total at three different flight speeds based on percentage speed potentials of the DJI Phantom 3 pro, according to the PIX4D capture app, at 70% (~5.3 m/s), 80% (~8.5 m/s) and 100% (~14.5 m/s) of the maximum capable speed. Flight altitude was set to 100 m for all flights.

Fixed-wing: Fixed-wing UAVs are limited by their stall speed. In addition, wind speed affects performance, and thus it was impractical to attempt to test at varying speeds. Therefore, the accuracy of the tags was only assessed at a single target groundspeed set to 18 m/s. Twelve tags were placed (7 scattered; 5 within a 1 m^2 area) in a 3.69 ha field, and GPS location referenced (MyGPSCoordinates

app; Kevin Willet, TappiApps) (reported accuracy +/-5m). A flight grid was created ~960 x 960 m (92 ha) in size, which included 19 lines at a spacing of 50 m, and flight altitude set to 100 m.

3.2.2.4 Attaching RFID tags to the sheep

Ethical approval was obtained for animal trials. The work described was conducted in accordance with the requirements of the UK Animals (Scientific Procedures) Act 1986 and with the approval of the IBERS Animal Welfare and Ethical Review Board. The work was also discussed with UK Home Office inspectors during routine visits. Thirteen Herdwick sheep (*Ovis aries*) were selected for tag application. Of the 13 sheep, two had a single tag attached to each of their horns, nine had a tag attached to one of their ears, and two had tags attached to collars fitted around their necks. In attaching tags to the horns, tags were first dipped in IMPACT adhesive glue (Bostik Ltd, Stafford, United Kingdom), placed on top of the horn facing skywards, wrapped with a crepe bandage, then secured with a layer of RHINO cross weave fabric tape on top (Ultratape House, Dundee, UK). Each ear tag was secured to the outside edge of an existing 'loop' management tag using two cable ties. When attached to collars the tags were cable tied to the back of the collar facing upwards. The sheep were held in the same 3.69 ha field where the fixed UAV accuracy testing was undertaken. A similar flight grid was created (~960 x 960 m) and flight altitude set to 100 m. A total of 7 missions were completed over a two-week period.

3.2.2.5 Data analysis

Duplicate hits (as a result of the tag sending three responses per transmission) that shared the same position were deleted. Any duplicate responses that occurred after GNSS update were treated as standalone responses as they had differing locations to the first response in the package. Only hits received along the grid lines were used, removing any that were recorded during launch/landing. Mean coordinates (Lat/Long) of each individual tag were subsequently calculated in open source GIS software (Qgis vers 2.12.3 *Lyon*). For assessing static accuracy, distance (in m) between each calculated tag mean coordinate (Lat2, Long2) and known GNSS location (Lat1, Long1) was completed in Microsoft Excel using the following formula, which is based on the Spherical Law of Cosines:

$$acos\big(sin(lat1) * sin(lat2) + cos(lat1) * cos(lat2) * cos(long2 - long1)\big) * 6371$$

UAV speed for each run was calculated using a custom-made script programmed in C++. The distance and time between each consecutive ear tag detection within a run was calculated, and speed consequently derived (m/s) for each observation. Subsequently, a mean speed (m/s) for each

individual tag ID was formulated.

Statistical analyses were performed using R Studio (R version 3.6.1) (Team 2013). The packages MASS (Bates *et al.*, 2014) and lme4 (Venables & Ripley, 2002) were required. The 95th percentile of the data was used as a measure of overall static accuracy.

Three regression analyses were performed. Firstly, in order to assess the relationship between accuracy (a positive, skewed, continuous variable), speed (continuous variable with values roughly close to 5 m/s, 8 m/s, 14 m/s and 18 m/s) and number of hits, gamma regression was used with speed and number of hits as explanatory variables. The functional form (i.e. whether higher order terms for hits was required) was motivated by local polynomial regression. Secondly, a (gamma) mixed effects model was used to provide an estimate of between-tag variability in accuracy. Thirdly, in order to assess the relationship between hits (an overdispersed count variable [mean = 21.84, variance = 400.82]) and speed, negative binomial regression was used. The Akaike information criterion (AIC) (comparing the model for hits predicted by speed, and a model for hits including only an intercept) was used to assess whether speed explained variability in the number of hits.

3.2.3 Results

The R95 parameter was calculated to be 58.5 m (n =175, M=29.6 m, SE=1.46) (Figure 3.4).

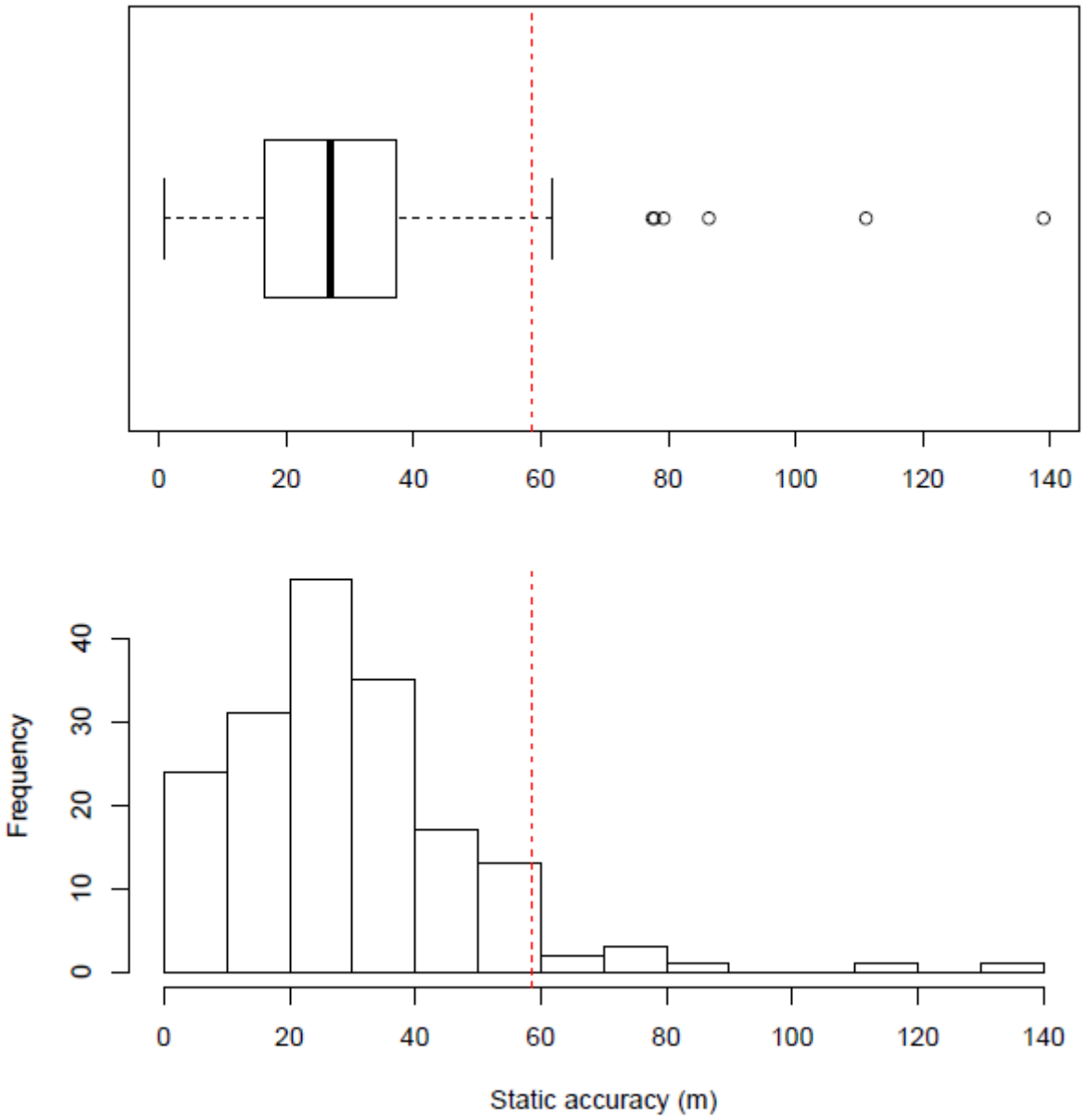


Figure 3.4 A boxplot and histogram of the static accuracy measurements, defined as the distance between each calculated tag mean coordinate produced from the RF system data, and known position of the RFID tags. The vertical red dotted line shows the R95 parameter with a value of 58.5 m.

Table 3.1 Gamma regression output with speed and number of hits as explanatory variables. Variable hits were scaled and centred to avoid co-linearity issues. Only terms for hits were statistically significant.

	Estimate	**SE**	***t* value**	***p* value**
(Intercept)	3.11	0.12	25.81	<0.001
Speed	0.01	0.01	0.91	0.36
Hits	-8.9×10^{-3}	3.9×10^{-3}	-2.29	0.02
Hits Squared	3.7×10^{-4}	8.1×10^{-5}	4.63	<0.001

The multivariate analysis showed hits to be a more important variable in determining accuracy than speed: terms for hits were statistically significant (Table 3.1), whereas speed was not statistically significant after accounting for effects of variation in hits. However, speed will influence the number of hits (Table 3.2): higher speeds tend to result in fewer hits, as shown in Figure 3.5. Figure 3.6 shows the observed relationship (as estimated by local polynomial regression) between static accuracy and number of hits, where lower values of accuracy imply better accuracy. Mean accuracy improves as the number of hits increases, but only up to around 40 hits. Thereafter, the mean accuracy declines for a larger number of hits. Beyond 40 hits, the limited data available means the estimated fit should be viewed with caution.

Table 3.2 AIC scores for the negative binomial models for the number of hits. Including variable speed vastly improves model fit, suggesting speed explains variability in the number of hits.

Model	**df**	**AIC**
Intercept and Speed	3	1349.57
Only Intercept	2	1433.14

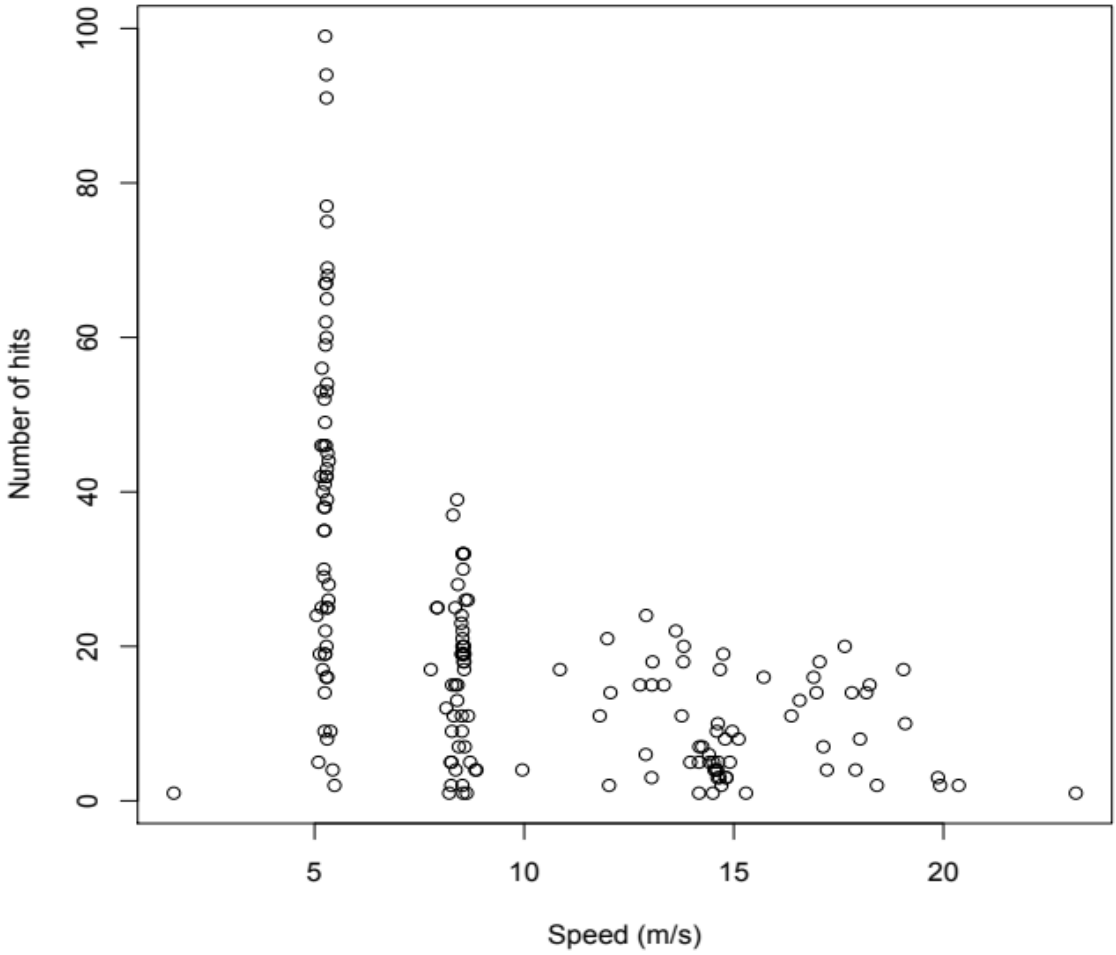


Figure 3.5 Effect of UAV speed (m/s) on observed number of hits (i.e. each successful package received from an RFID tag) by the RF tracking system.

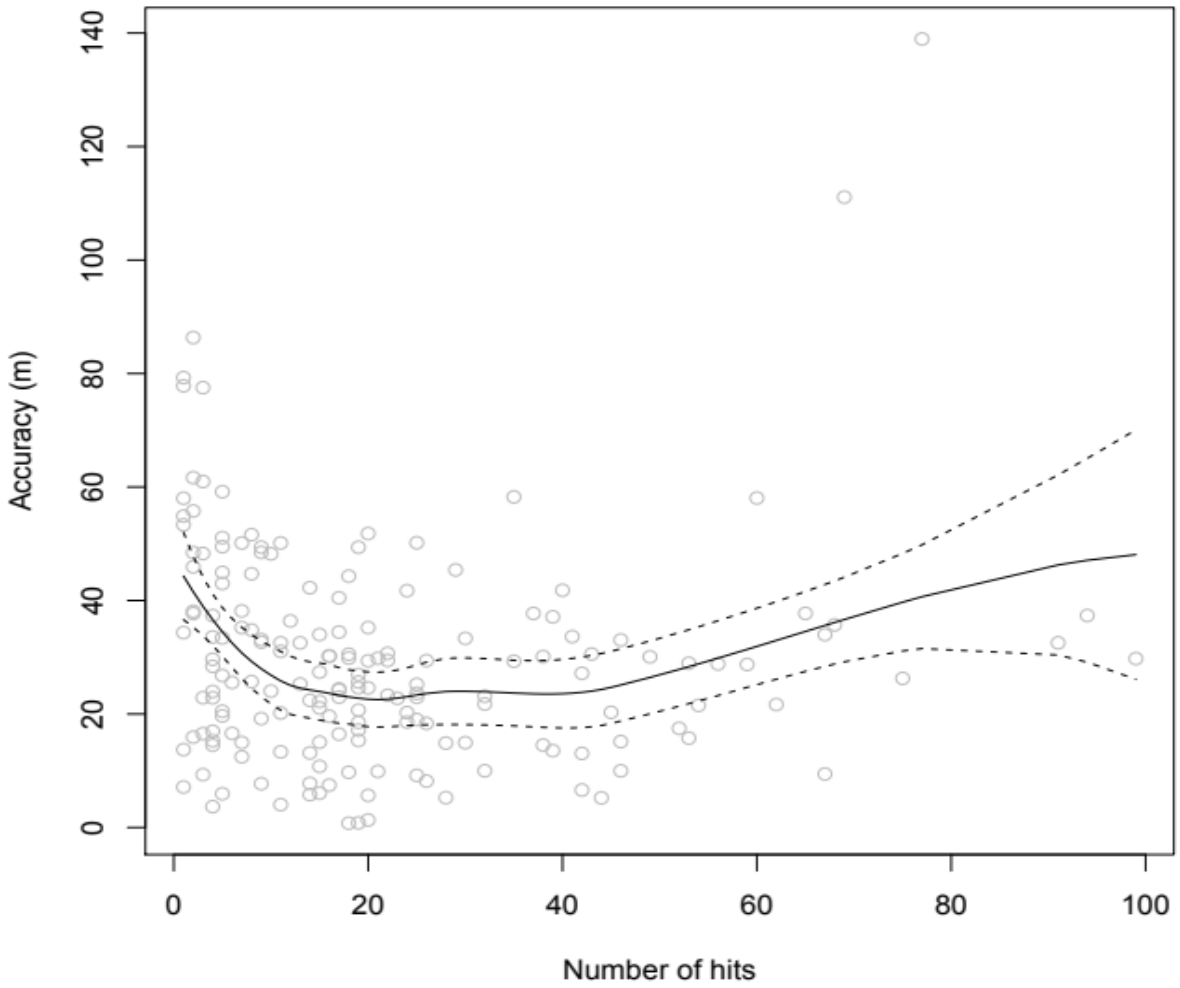


Figure 3.6 Effect of number of hits (i.e. each successful package received from an RFID tag) on observed static accuracy (i.e. distance between each calculated tag mean coordinate produced from the RF system data, and known position of the RFID tags) of the 175 data points (m). The black line shows the observed relationship between the mean accuracy and number of hits as estimated by local polynomial regression, and the dotted lines are 95% confidence intervals. Between 10 and 40 hits generally gave acceptable accuracy. Beyond 40 hits, the limited data available means the estimated fit should be viewed with caution.

All the RFID tags attached to the management ear tags of the sheep (n=9) and one attached to the collars were still in place at the end of the experiment (n=2), however none of those attached to the horns (n = 2) remained. The tags that fell off did so before the first flight had been undertaken. Of the 10 still attached to the sheep, 9 worked with 100% consistency across all 7 recorded missions, while one failed after two missions. Tag response reliability was therefore calculated as 93% (100/70*65).

3.2.4 Discussion

3.2.4.1 System accuracy

It was reasoned that the variance witnessed in the data is likely due to component variation and suspected temperature compensation issues with the RF components, which remain to be fully quantified in future work. That said, including a random effect to assess variability between tags, the estimated between-ID tag variance was 0, indicating no between tag variability (i.e. no individual tag was inherently more accurate than another). Overall, the comparatively high level of accuracy achieved as a low-cost RF based system likely outweighs the observed variation in precision. Though further work is required, the results indicate that in considering static accuracy, increasing the number of hits (either by decreasing the time between pings transmitted on the tag trigger/or decreasing UAV speed) are key factors worth exploring.

3.2.4.2 Considerations and critiques for using the system

Since positioning is calculated from first to last response of each individual tag within the flight, only a time period and not a precise time point can be ascribed to calculated position, with the length of the time period depending on the size of the flight grid being performed. Used in conjunction with GNSS systems (e.g. tracking collars), clarity could be improved. For example, GNSS tracking of a small number of animals would provide a high number of consistent recordings over a time period, whereas the UAVRTS could deliver a lower temporal number of 'snapshot' recordings for the entire group of animals. A critique of the system could be that flights would only likely occur in fair/moderate weather conditions, and therefore recordings would not represent distribution under all meteorological circumstances. Therefore, having GNSS trackers present at the same time would also help by providing data for instances where the UAVRTS could not.

When preparing flight grids, specifics such as line width are comparatively minor considerations relative to the need for grid to be large enough, and the tag trigger power to be low enough for the UAV to be able to fly out of range of the tags. A criticism of this system could be that a rough radius of all combined tags in an area must be known in order to construct a grid to cover them all. When used on ungulates (who typically herd), or in conjunction with GNSS trackers already on the ground, this may be more easily definable. Furthermore, the use of quick reconnaissance flights could be used to identify the spread of the target group. Another limitation is that although fixed-wing UAVs are capable of a large range the maximum grid size may be limited by the UAV operating regulations of the respective country. Though beyond visual line of sight (BVLOS) authorisation has the potential to extend the operational capability, this is still a developing framework in many countries and therefore may not be immediately accessible. Under current circumstances flying adjacent grids of a legal size

in succession is a workable alternative. Finally, concern has arisen in the literature with regard to the use of UAVs in animal studies, and the effect this may have on animal behaviour. A variety of experiments have assessed the effect of UAVs on a range of target species (Ditmer *et al.* 2015; Vas *et al.* 2015; McEvoy, Hall & McDonald 2016; Rümmler *et al.* 2018). The consensus from these papers is that whether a UAV will disturb the target animals will depend on many factors (e.g. operational altitude, survey design, UAV design, target species). In the majority of publications, there were no behavioural response on their target species above low altitude (~30-50m). In one study (Vas *et al.* 2015), the authors even remarked about the ability to fly their UAV as close as 4 m from the birds without any detectable behavioural response. With the regard to long term exposure concerns, (Rümmler *et al.* 2018) actually observed evidence of habituation to UAVs. Within the experience of conducting UAV flights in this study, no obvious change in behaviour was witnessed at any point as a result of the UAV. It can be agreed that observing good UAV practice and awareness of the target's behaviour when using UAVs is important and necessary to avoid any potential disturbance.

3.2.4.3 Further research

Though the viability of this system across multiple UAV platforms has been demonstrated with considerations for its use provided, several key issues would benefit from further research and development. The system's performance in situations where animals may be in shaded/covered locations (e.g. woodlands, rocky areas) needs to be investigated before it is deployed in such circumstances. In addition, integration of the RF system with the UAV autopilot modules would allow more sophisticated surveying methods, such as circling or slowing down when a tag is detected in order to increase accuracy further.

3.2.4.4 Conclusions for the developed UAVRTS

This study presents the first-cross platform compatible UAVRTS. Its flexibility and low-cost nature, together with the degree of accuracy achievable and proven ability to be utilised on mammals demonstrate its readiness as a field ready tool. Though applicable in many environments/situations, in the larger context of movement ecology, the current suitable applications of this system would be: 1) the tracking of large ungulate herds; or 2) target animals which are located in enclosures or have defined home range areas.

3.2.5 Using the UAVRTS within the current PhD study

While the validation of the UAVRTS technology was ultimately successful, the reality was that it took the first 24 months of the study to complete. By which point most of the limited budget in the study had been exhausted, as well as a considerable amount of time. As such, large-scale tracking of all the animals on the sites was not possible within the current study lifetime. As it became apparent that the developed UAVRTS solution could not deliver the animal location data required to achieve the main PhD and chapter objectives, additional options had to be explored.

Though a number of issues with GNSS tracking were outlined in the chapter introduction, it was decided that tailoring GNSS hardware solutions to the required experimental design of the study could prove effective in capturing the data required for completing the biological objectives. It also could result in novel approaches to both implementation and data handling, which could be beneficial to the wider animal tracking literature.

3.3 Collecting GNSS data efficiently, and processing the data

3.3.1 Global navigation satellite systems

3.3.1.1 The underutilisation of GNSS in livestock tracking studies

Over the past 25 years, global navigation satellite systems (GNSS) (e.g. GPS, GLONASS, Galileo, BeiDou) have become a key tool for scientists wishing to study distribution of animals. The resolution attainable, both spatially and temporally make it the most utilised method for studying ungulate movement (Anderson, Estell & Cibils 2013). However, making optimal use of this technology is not often achieved, as researchers and conservationists alike regularly struggle to maximise the operational capability of the tracking devices given two chief constraints of GNSS collars already outlined: hardware cost, and battery life.

3.3.1.2 The rise of low cost GNSS loggers

As alluded to previously, high individual unit costs of tracking units often correlates with low numbers of animals being tracked (Allan et al. 2013; McGranahan, Geaumont & Spiess 2018), and thereby the positional information of only a few individuals being used to represent whole herd/group movements (Anderson, Estell & Cibils 2013). Given the high price of many commercially available GNSS units, not only has their use been mainly restricted to research studies on a low number of sampled individuals, but their potential as monitoring devices for land managers, used with either wildlife or livestock, has also largely been unrealised. The unit price of commercial tracking collars (~£500-1500 depending on specifications) often exceeds that of the subject animal itself many times over (particularly when tracking livestock). As a result, in recent years increasing interest has been given to the development of open source Arduino based GNSS loggers (Cain & Cross 2018; McGranahan, Geaumont & Spiess 2018), with the goal of making them low-cost and customisable. Detailed information on their construction is available, and with low reported hardware costs of $40 (Cain & Cross 2018) and $125 USD (£32 and £100 at the time of writing) (McGranahan, Geaumont & Spiess 2018) reported respectively, and with a static accuracy (95th percentile) of 4.0 m (based on logged positions at 20-s intervals for 90 min) (McGranahan, Geaumont & Spiess 2018), these loggers present an attractive option for scientists and practitioners going forward.

3.3.1.3 Balancing battery life with the data required through mixed frequency scheduling

The second constraint of using GNSS loggers is battery life, particularly in longer term studies (e.g. wildlife studies) where regular collar retrieval (for battery recharging) is not possible. The key implication of this is the effect on the sampling frequency used for GNSS recordings, the simple notion being that the longer the dormant intervals between recordings, the more battery life conserved and therefore the longer the study may persist.

Within GNSS tracking studies, the option of scheduling for different recording frequencies within the same experimental session is not one that is typically exploited, with studies often only using a set interval recording rate throughout. Given that higher recording rates consume more battery life, the interval length is often dictated by how often the loggers can be retrieved (so that batteries may be recharged or replaced), with recording intervals of 5 min through to several hours being commonly used. However, as the length of the interval between recordings increases there is typically a corresponding decrease in the accuracy of GNSS tracking units themselves (Swain, Wark & Bishop-Hurley 2008). This is primarily due to the configuration of the tracking units. Given GNSS units are put into a low power state when not recording, when they do activate for a new recording session, they require an amount of time to search for available satellites, and then refine their position. However, units often have a set amount of time ('window') in which to do this (90-240 secs often) before switching off again to save battery. If the last recording session was relatively recent (e.g. a few mins to a few hours before), the position of the satellites will not have changed drastically and therefore the 'time to first fix' (TTFF) will be quick, which leaves more time for the refinement of the GNSS position, and therefore better accuracy. This is termed a 'warm or hot' start (depending on the exact interval length between recordings). If, however the interval length is great (e.g. hours or days), then it will take longer to search for possible satellites and collect information from them ('almanac'). This will therefore result in a slower TTFF, and therefore the refinement period will be shorter, which will typically lead to less accurate positioning. Though lengthening the GNSS window time would lead to more accurate positioning even for longer interval lengths, this obviously decreases the battery life. Therefore, having a GNSS window time which delivers reasonable accuracy with modest battery consumption is usually the opted compromise. Mixed frequency scheduling therefore offers another advantage. If, for example a high frequency recording rate of 1 position per second was recorded for 5 min, then the total GNSS window for this time would be ~five min. This would obviously consume more battery, but if placed at the beginning of a cold start session, this would have the benefit of ensuring that sufficient accuracy is achieved whilst maximising the amount of data collected. The use of a mixed frequency recording schedule of low and high intensity resolution data can therefore offer a useful compromise between extending battery life and maintaining accuracy. Different frequency data can also be used for different analyses (Figure 3.7) which when used effectively can complement each other. For example, low temporal resolution spatial data over a given period can be used to define the overall area utilised by a given individual, whilst embedded periods of intensive recordings deliver the required resolution to infer the behavioural aspects of an animal within the same sampling period (Rutter 2007).

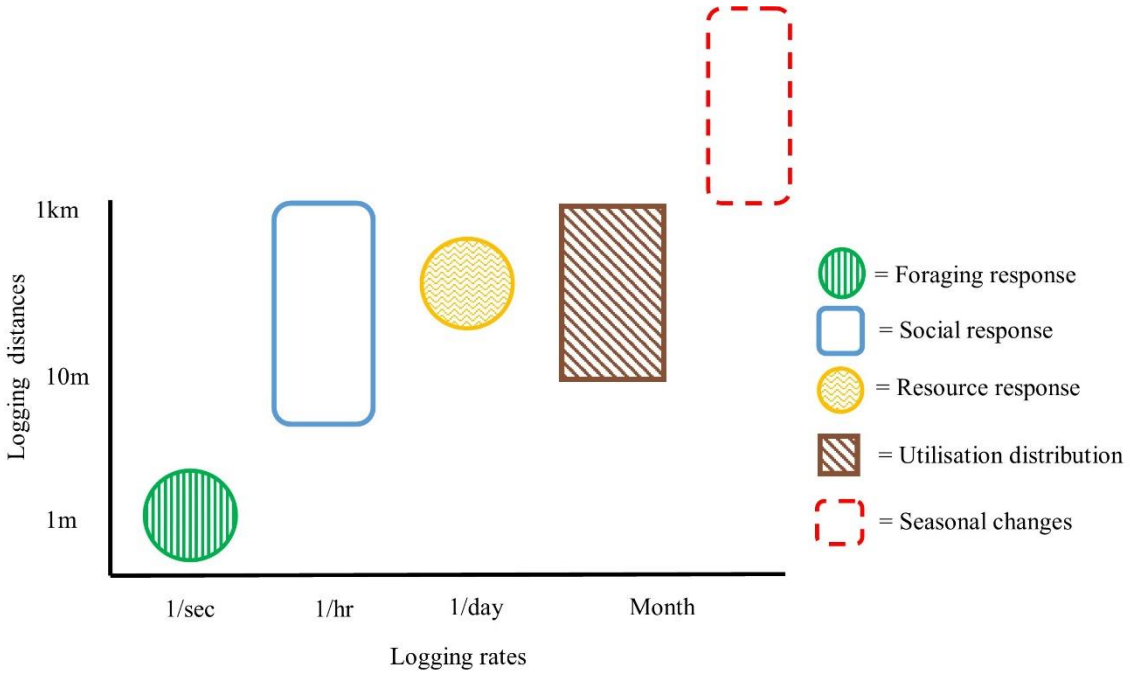


Figure 3.7 Spatiotemporal scales of different frequency recordings and the respective insights they can provide into movement ecology of an animal.

3.3.1.4 Why classify animal behaviour?

The behavioural classification of animal movement (e.g. foraging, moving, resting) can be used to measure the effect of different physiological (e.g. parasite burden or parturition) and environmental pressures (e.g. climatic conditions) on individual animals (Dobos *et al.* 2015). It can present a different angle to other analytical options by allowing the response of an animal to its environment to be considered rather than vice versa. By its very nature, behavioural classification is also not typically burdened by the concern of autocorrelation which makes it ideal for high frequency data recording. Though studies often use 5, 10, or even 20 min interval data (Ungar *et al.* 2005; Putfarken *et al.* 2008; Swain, Wark & Bishop-Hurley 2008; Dobos *et al.* 2015), using 1s interval data is ultimately best as it does not assume a movement pattern between recordings. This is because animal movements between position samples almost always rely on a straight-line estimation. This therefore creates the assumption that animals are moving in linear fashion between recorded points, which means that the longer the intervals between recordings, the greater the likelihood of error in the behaviour inferred. Secondly, given GNSS window lengths for a single fix are often 90-240s and therefore the units are activated and consuming power anyway, collecting as higher frequency data as possible is an efficient way of maximising the amount of data collected for battery used.

3.3.1.5 Why complex models are not suitable for many situations

Inferring animal behaviour can often be done on a visual basis (Torres *et al.* 2017), especially when viewed with environmental layers (e.g. basemaps, vegetation classifications) which can help inform

on any assigned behaviour by providing context to the movement area. However, this can be subject to human bias, is labour intensive, and is therefore only realistically suited towards small datasets (either containing few animals, or short sampling durations), or as a ground-truthing tool. Automated behavioural classifications methods are therefore preferable for most data sets. However, many of the quantitative methods of inferring behaviour in use (e.g. Bayesian state-space models, hidden Markov models) exist primarily as research tools, and are often statistically complex, and require a good level of programming ability. Without the necessary specialist skill and knowledge, adopting one of these methods can be prohibitively challenging. As mentioned previously, in the case of livestock-based tracking, many studies are undertaken in applied situations and by conservationists, land managers etc., and therefore the use of such classification methods are often restrictive.

An alternative approach of using animal velocity as a metric to help discern behaviour is one which has been well adopted (Putfarken *et al.* 2008; Dobos *et al.* 2015). Whilst arguably cruder than other methods (as they do not assess behavioural patterns, but simply use velocity), it has often proved to be accurate and succinct in the studies where it has been employed. Furthermore, since this method lacks strong transferability between taxa, given that it requires prior knowledge of the subject animal's behavioural status at different velocities, this approach ensures that appropriate ground-truthing needs to take place in cases when a new type of animal is being studied. Though initially more practically laborious, this adds a level of rigour to the analysis compared to other methods that may not so obviously require taxon-specific information.

3.3.1.6 Developing a GNSS tracking solution and processing tools

In summing up the points considered in this part, it was considered that using low-cost loggers, either from an open source design or through self-build construction was the most suitable way of maximising the number of individuals tracked, and best way to improve sample representation. Secondly, it was determined that using mixed frequency scheduling would be the best compromise between the need to conserve battery life whilst collecting sufficient data to provide meaningful outputs from analyses undertaken. Finally, it was concluded that many of analyses currently available for inferring animal behaviour are complex in their use, and that using a velocity based metric approach would be sufficient for this study and others where land practitioners would seek to utilise it.

With all this considered, a number of experimental objectives were created in order to best reflect these conclusions within the GNSS tracking operations of the study. The first objective was to produce low-cost GNSS loggers which would have mixed frequency scheduling capability and design a schedule that could last the whole tracking sessions between handling opportunities on both sites (~50 days). The second objective was to create software that could automate and simplify the data handling from these loggers: 1) by separating different temporal resolution data within a single data

set; and 2) by infer resting, foraging and moving behaviour from the subsequent high-resolution data generated. The software would need to be quick and simple to use, and require no previous programming ability, with detailed instructions provided for its use.

3.3.2 Logger design and deployment results

3.3.2.1 Logger design

The GNSS loggers used were custom-built in order to provide low power, schedulable logging at a low cost. They were built around the ATMega328P microcontroller on a custom PCB that used a Quectel L86 GNSS module and a low-cost Sparkfun OpenLog micro-SD data logger. The microcontroller was programmed to read a schedule from a plain text file on the micro-SD card and to subsequently collect data at the intervals specified. The data was saved as plain text to the micro-SD card and could then subsequently be accessed by the user from any conventional computer using a micro-SD card reader. The schedule file allowed for 3 different patterns to be used simultaneously. For example, the first schedule might collect 1 min of data every hour, the second might collect 1 min of data every 10 min between 6 am and 6 pm, and the third might collect data continuously for 1 h on every 10th day of deployment. The loggers were powered from a rechargeable 1500mAh lithium polymer battery and were housed in a rugged plastic box that was attached to a flexible collar. When logging over extended periods the power consumption was approximately 0.015mAh per log entry, which when combined with a small continuous background current results in approximately 50,000 log entries from a 6-week deployment.

3.3.2.2 Test environment

Tracking operations were undertaken at the two upland moorland sites in mid Wales: Ffridd Fawr at RSPB Lake Vyrnwy (73 ha, 52.797790, - 3.576938) and Penglaneinon hill in the Elan Valley (111 ha, 52.253826, - 3.614791), between May and July 2018. Ten Welsh Mountain sheep were collared, and subsequently tracked at each site, resulting in a total of 20 tracked individuals. Given that the rate of travel boundaries for different behavioural classifications had already been validated for sheep and proven to be of high accuracy (87%) by Putfarken *et al* (2008), and further tested by Dobos *et al* (2015), their boundary values were deemed suitable for this study. Due to the size of the sites and the difficulty of the terrain, practical necessity dictated that handling of the sheep took place only when existing management operations occurred, i.e. collar deployment when sheep were placed on the sites and removal during shearing time. The length of the corresponding tracking sessions was 45 d at Ffridd Fawr and 49 d at Penglaneinon. The schedule implemented on the loggers was similar for all tracked animals at both sites: hourly recordings for 7 d, followed by an eighth intensive day where 1 s recordings were taken for the first 5 min of each hour, followed by recordings once every 5 min for the rest of the hour. The configuration of mixed frequency recordings of high (1 s), medium (5 min),

and low (1 h) resolution data was chosen to explore the utility of each frequency level, as well as ensure sufficient battery life for the whole recording session.

3.3.2.3 Logger deployment results

Of the 20 collars deployed, 11 were still recording data on removal. Of the nine no longer recording, 7 loggers contained partially recorded data ranging from 1 to 45 d, whilst 2 loggers failed to record any data. Six of these loggers were found with either dislodged SD cards or batteries, a further two were thought to have suffered malfunctions during hardware assembly, and a final collar suffered water damage due to rainwater ingress on a faulty logger housing. Despite the mixed success of the logger assembly (which was anticipated given their experimental nature), the battery consumption performance was as anticipated.

The data was found to be correctly recorded according to the specified schedule, meaning a total of 6 intensive recording days at both sites as well as 42 and 46 d of 1 h interval recordings at Penglaneinon and Ffridd Fawr, respectively. However, it was noted that a number of the recordings in both the hourly and 1 s interval fixes contained ‘frozen’ coordinates (i.e. the logger had failed to refresh the location coordinates from one recording to the next). In the case of the hourly recordings this was found to be due to an error in the logger firmware which caused a failure in location updating. This was due to the way in which the firmware kept track of the age of the last acquired GNSS location. The logger used the microcontroller's clock to determine the time since the location was updated, and only caused a refresh when it was more than 3000ms old. When the microcontroller was asleep between logging events this caused the clock to stop running and therefore the GNSS data was deemed to be recent even when several hours old. This flaw was only an issue when the logger was running single point schedules for extended periods. Frozen coordinates in the 1 s recordings were found to have occurred due to a default speed threshold in the firmware of the GNSS module used. Under this function, new recordings were ignored if their distance from the previous recorded coordinates was small (<5m), thereby causing the modules to default to previous values. The intention of this function is to minimise likely GNSS noise when the logger is in a static position, however it is unsuited in studies such as this where the behavioural classification relies on the detection of small movements.

In the case of both the hardware and firmware flaws witnessed in this study, it can be argued that extensive pre-testing before the main deployment would identify many of errors associated with logger deployment in this study. However, these errors typify those often observed when using custom-built open-source loggers, where in reality the level of pre-testing undertaken will more often than not still result in undetected flaws being present when main deployment takes place. It is therefore essential that users are vigilant when viewing the raw data for any inconsistences which could indicate irregular recordings. However, given the number of recoding points for a single logger

in a tracking session can be in 10's of thousands or even millions (indeed, a complete recording sequence for the first Penglaneinon session was ~45,000 recordings) it is therefore imperative to include filters and quality checks within the data processing software used.

The need to be certain of the data integrity was a key motivation for developing data processing tools in this study. It became clear that the checks for data consistency would exist in the form of a series of customisable parameters, which should have the capacity to be easily altered by a non-programmer. Given the difficulties in identifying easy to use behavioural inference tools within the literature, it was decided that the behavioural analysis of the high frequency data could be integrated with the data processing tools within the software to allow quick, complete processing flow from raw data to processed outputs. This was the basis for the creation of Beferrer (an amalgamation of behavioural inference).

3.3.3 Beferrer design, functions, and parameters configurable

The Beferrer application (https://github.com/NealSnooke/Beferrer.git) created is a cross-platform JAVA-based application that runs on directly on the command line of recent versions of Windows, Linux, and Mac OS. The application is split into two primary functions. Firstly, it imports raw tracking data, removing any erroneous lines containing non-standard data or illegal characters. It sorts the data according to the mixed frequency recordings (e.g. 1 s, 1 h) present. The application accepts comma delimited text files as input, with each valid line containing four items: latitude, longitude, date, time. Latitude and longitude are decimals. Dates are in dd/mm/yyyy format and times in hh: mm: ss format. An example of a valid line of data is: *52.439037, - 3.993740, 27/5/2018, 10:5:54*. Any lines that do not contain four items in this format are ignored.

This first function was also compiled into a separate second programme called 'Schedsplitter' (an amalgamation of Schedule splitter), so that users wishing simply to separate the different frequency data for other analyses could do so without needing to produce the behavioural inference results.

The second function is to then separate data by chosen interval intensity into named behavioural classification classes (e.g. moving, foraging, and resting). The entire pipeline is customisable with 18 input parameters configurable. Broadly these are summarised in the following (See appendix A.3 for full parameter list).

3.3.3.1 Activity

Firstly, the considered time resolution of the data for analysis may be configured (-g,--GroupTime <min:max>). Each sequential set of coordinates that satisfies a stated range is considered a sample group, i.e. a track segment that includes movements between min and max secs duration. This allows sample group sequences to be extracted from mixed resolution data sets. Following this, the distance

between each set of coordinates and the next are calculated using the Haversine Formula, velocities (m/s) are then determined, and behavioural states classed according to pre-defined velocity rates. Though initially designed for livestock, the application is non-taxon specific. Any specific rate of travel values, where known for different behavioural classifications of an animal, may be inputted (-a,--Activity <b1,s1,b2,s2...>). Finally, though a behavioural class is allocated between each sequential set of coordinates, it is also possible to average the movement speed over a specified period of time (-m,--MovementAverage <time>). This feature is useful for users wishing to define the length of behavioural periods most suitable for their specific study. For example, in assessing grazing areas, where short resting/moving phases between grazing segments are often included in grazing periods

3.3.3.2 GNSS error/irregularities

Beferrer also features a number of parameters to help limit potential GNSS logger error and irregularities in their recordings. This primarily corresponds to the accuracy of recordings relative to their true location, and instances where readings are not recorded properly. Given that the velocity values of the differing behavioural classes are separated by such short distances (<1 m per m/s), and the reported standard accuracy for GNSS loggers without differential correction is typically greater than this (~10 m, though dependent on sampling rate), it is impossible to be completely certain that the distances between sequential set of coordinates are a result of animal movement and not GNSS error. However, in reality GNSS error in this study was most often observed as either single haphazard recordings which deviated greatly from the previous recordings, or as a whole series of recordings which were off-centre from a true movement track. The actual general structure of the tracks routinely remained true, and therefore confidence in the behavioural classes reflecting true animal movements could be had.

Within Beferrer it is possible to set a velocity threshold, whereby a data point will be removed if movement within adjacent points is above the specified value (-cg,--GlitchSpeed <speed>). This will however only remove a single point, and in the event of multiple erroneous high velocity recordings, a second parameter may be utilised whereby all movements before or after a movement greater than a specified speed are removed from each group (-gg,--TrimGroupBelowSpeed <speed>). The velocity threshold used may be derived from a known top speed of the study animal, or from a velocity where beyond which non-typical behaviour would likely ensue, and thereby may not be considered useful for the behavioural classifications being inferred. Assessing the velocity of a given animal across a recording session is one way of deciding this. For example in Table 3.3, for this study it was proposed that a speed threshold which was exceeded < 1% of the time to be a suitable cut-off point for minimising the likely encroachment of GNSS error in the data (4 m/s velocity threshold).

Table 3.3 Frequency of recordings in relation to velocity from the moving data of 10 collared sheep at Penglaneinon (23/05/18-11/07/18). Velocities were calculated using 1-2 sec interval data. All stationary recordings were removed so only moving points were considered

Velocity (m/s)	Frequency	Percentage of total frequency (%)	Accumulated percentage (%)
0-1	136210	86.97957	86.97957
1-2	14861	9.489783	96.46935
2-3	3243	2.070881	98.54023
3-4	1250	0.798212	99.33845
4-5	632	0.403576	99.74202
5-6	189	0.12069	99.86271
6-7	111	0.070881	99.93359
7-8	30	0.019157	99.95275
8-9	28	0.01788	99.97063
9-10	22	0.014049	99.98468
10-11	8	0.005109	99.98979
11-12	8	0.005109	99.9949
12-13	4	0.002554	99.99745
13-14	1	0.000639	99.99809
14-15	3	0.001916	100

It was also decided that optional parameters be included which can help mitigate against erroneous recordings that may be caused by firmware errors, such as frozen coordinates (as in the case of this study). That said, the complication of these frozen recordings for inferring behaviour is that if it is suspected that a dataset contains recordings likely due to logger error (e.g. abnormally high proportion of resting periods, or resting periods extending for unusually long times for the study animal), it cannot be completely ascertained which exact stationary recordings are a result of error, or just an animal resting. The application does feature a parameter (-p,--PositionAverage <dist>) that averages position coordinates over a specified distance prior to generating movements. This has the effect of smoothing tracks in the presence of frozen points or even GNSS noise. When set to zero, this results in the combining of non-moving sequences into a single longer duration move. Nevertheless, users should be cautious in inferring behaviour in such situations, with clear consideration given to the appropriateness of using the data being based on the given circumstances.

Finally, in the scenario where animal movement is restricted to a physical boundary (e.g. fences) or animals have moved outside of the study area of interest, recordings outside a specified zone may be ignored (-c,--CoordinatesFence <N,W,S,E>).

3.3.3.3 Temporal focus

Beferrer also allows specific temporal analysis, which can be useful for users wishing to study behaviours at particular times of day, or for tracking behaviours across specified time periods (e.g. weeks, months). Designated time periods during a day are set by two parameters. The first ignores daily data before a specified time period (-cts,--StartTime <hh:mm:ss>), and the second (-cte,--EndTime <hh:mm:ss>) after a specified time period. Recordings before a specified date may also be ignored (-cds,--StartDate <dd/mm/yyyy>), as well as after (-cde,--EndDate <dd/mm/yyyy>). Both date/time parameters may be used concurrently.

3.3.3.4 Ancillary

Finally, several parameters exist to tailor the input of files into Beferrer and assist in its use. Individual text files may be inputted (-f,--File <name>), or a input folder/directory may be specified (-fd,--Folder <folder>) where multiple files maybe processed. Files within a folder will be selected according to the specified file extension type (-fs, - -Suffix <suffix>) (e.g. .txt, .csv). The full list of parameters can be printed on the command line to allow easy editing of commands (-h,--help), and detailed output of the processing steps can be provided for explanation, or debugging (-v,--Verbose) (though a summary of the processing steps is outputted anyway to the terminal).

3.3.4 Beferrer outputs, and recommendations

Output files are created in the same folder as the input data. These file names include a compact version of the (non-default) command line options to allow output to be easily preserved if the tool is run with different options. The main outputs include a delimited text file (.csv) for each behavioural class, as well as a combined file will all the classes included. Within each file; location information (lat, lon coordinates) between each sequential recordings, raw and averaged processing values used for inferring behaviour (duration (s), distance (m), velocity (m/s), average velocity, as well as the designated behavioural class are recorded. In addition, Beferrer also creates an optional data file for the widely used (free) GNUPLOT tool (http://www.gnuplot.info/) to allow quick sense checking of the results, particularly with regard to potential GNSS error. Assuming that GNUPLOT is correctly installed, the contents of the file prefixed 'gnuplot_commands' can be pasted onto the command line (after setting the working folder to the location of the output) to produce a .eps graph. This file contains the gnuplot commands and the relevant data, in plain text format and can be edited as desired, e.g. to change line widths, axis ranges, titles etc. Each time Beferrer is utilised, the chart may be produced using GNUPLOT.

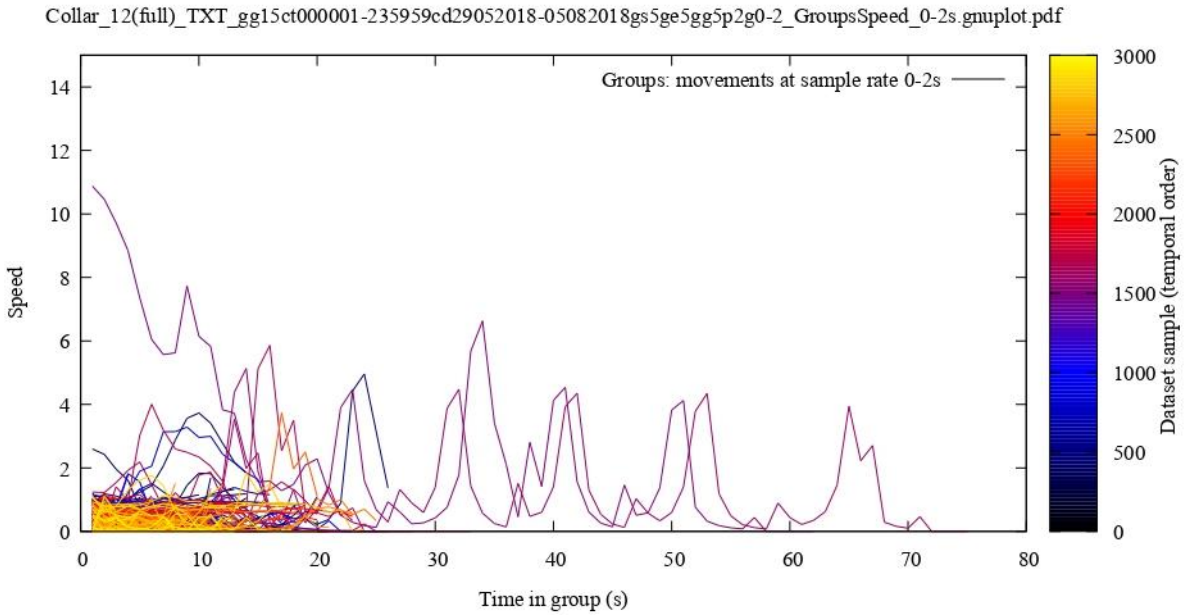


Figure 3.8 GNUPLOT output from Beferrer demonstrating animal speed within each sample sequence throughout the recording period for collar 12 using interval data at 1-2 secs. GlitchSpeed (-cg), set to 15 m/s, and TrimGroupBelowSpeed (gg) to 15 m/s.

The chart above (Figure 3.8) demonstrates the speed of the animal during each group sequence through the dataset. It is particularly useful in highlighting potential GNSS error/other non-standard movement. Short sharp peaks most often represent single or momentary glitches caused by GNSS error which quickly normalise as the group time goes on. Longer, more smoothed peaks will more likely be due to consistent GNSS error in the recordings or even nonstandard behaviour on the part of the animal e.g. a flight response caused by the animal getting startled. In all cases, users may wish to mitigate the observed error/non-standard behaviour, and therefore can rerun the Beferrer application with parameters reconfigured (GlitchSpeed (-cg), and TrimGroupBelowSpeed (gg)) towards a lower velocity threshold (Figure 3.9).

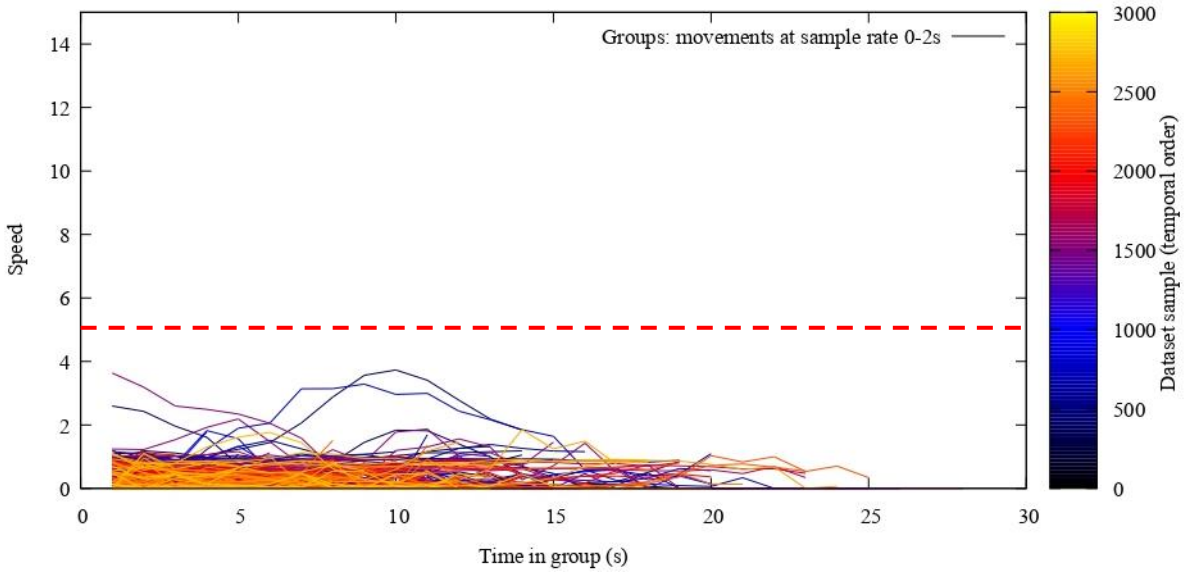


Figure 3.9 GNUPLOT output from Beferrer demonstrating animal speed within each sample sequence throughout the recording period for collar 12 using interval data at 1-2 secs. GlitchSpeed (-cg), set to 5 m/s, and TrimGroupBelowSpeed (gg) to 5 m/s (denoted by the red dashed line).

The text file outputs of Beferrer may also be inputted directly into Geographical Information System software (GIS), which when viewed with other environmental layers, can be useful in further ground truthing the inferred behaviour points. Figure 3.10 exemplifies a processed sequence consisting of 1 s interval data. In this instance, an individual sheep is inferred to be fluctuating between walking and grazing along a stream whilst stopping twice at the waterside before continuing. Intuitive observation would conclude that the individual is browsing along the stream edge for a suitable place to uptake water. This appears more evident when closer examination reveals the animal to be near-stationary for 12 and 20 s respectively at the waterside edge. Though this type of analysis on its own can never provide absolute certainty, it does provide a degree of confidence in the output either way and is easy and quick to undertake.

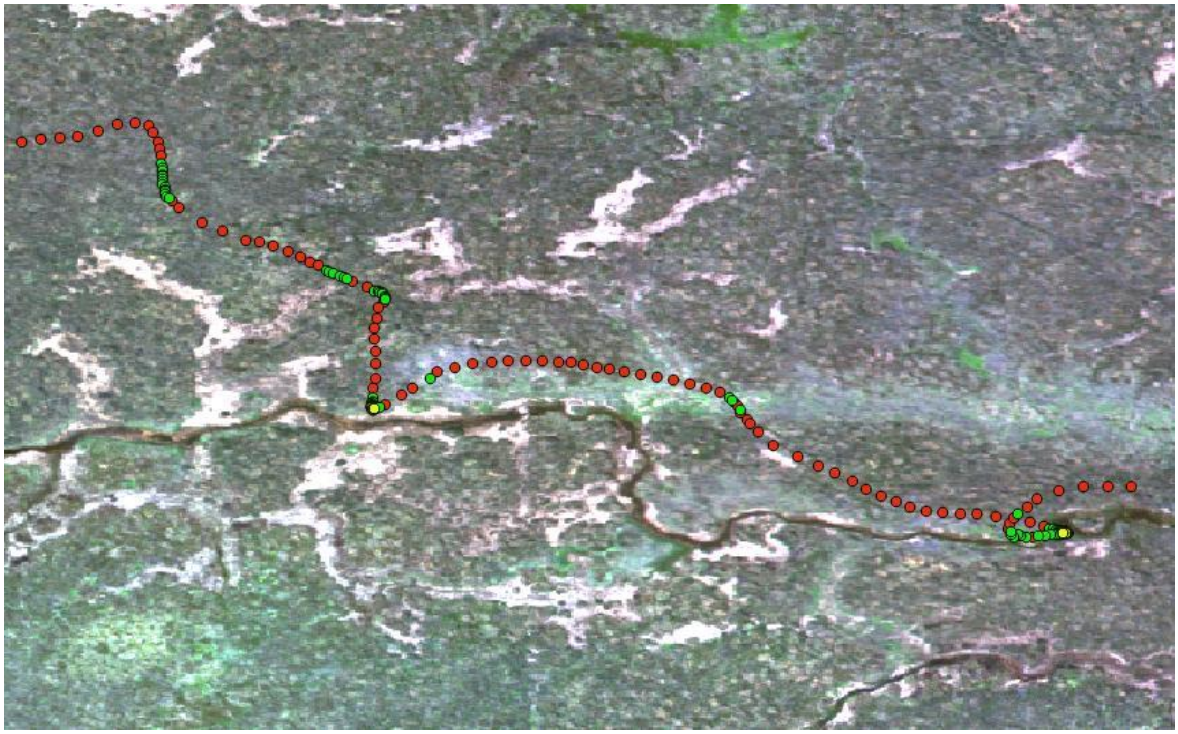


Figure 3.10 Combined behavioural classes from a single 213 s sample sequence of 1-2 s interval data from collar 12. Behavioural classes are; Moving (Red), Grazing (Green) and Resting (Yellow). Baselayer = Mulitspectral orthmosaic from a Micasense Rededge camera mounted on an autonomous Skywalker X8 fixed wing UAV.

GIS visualisation can also be used to compare the differences between raw (Figure 3.11) and averaged movement periods (Figure 3.12), and allow the user to form a judgement on whether it best to average the movement sequences or not, and if so which length of time might be suitable. In this example there appears to be little difference. The choice of whether to average the movement data will ultimately depend on the behavioural definitions used. For example, the concept of a 'grazing period' is subject to debate. It could be argued to include just the period where an animal has its mouth to the ground and is eating, however this would omit the premise of 'browsing', whereby an animal may walk within a grazed patch to select what to eat next. In truth, these definitions need not be fixed, and therefore as long as they are defined within the individual study, the movement average which best addresses the exact hypothesis concerned may be selected.

Given the speed by which Beferrer may be operated, repeat iterations of application may be undertaken until a user is satisfied with the observed likely level of error as well the conformity of the spatial distribution, and preferred temporal scale.

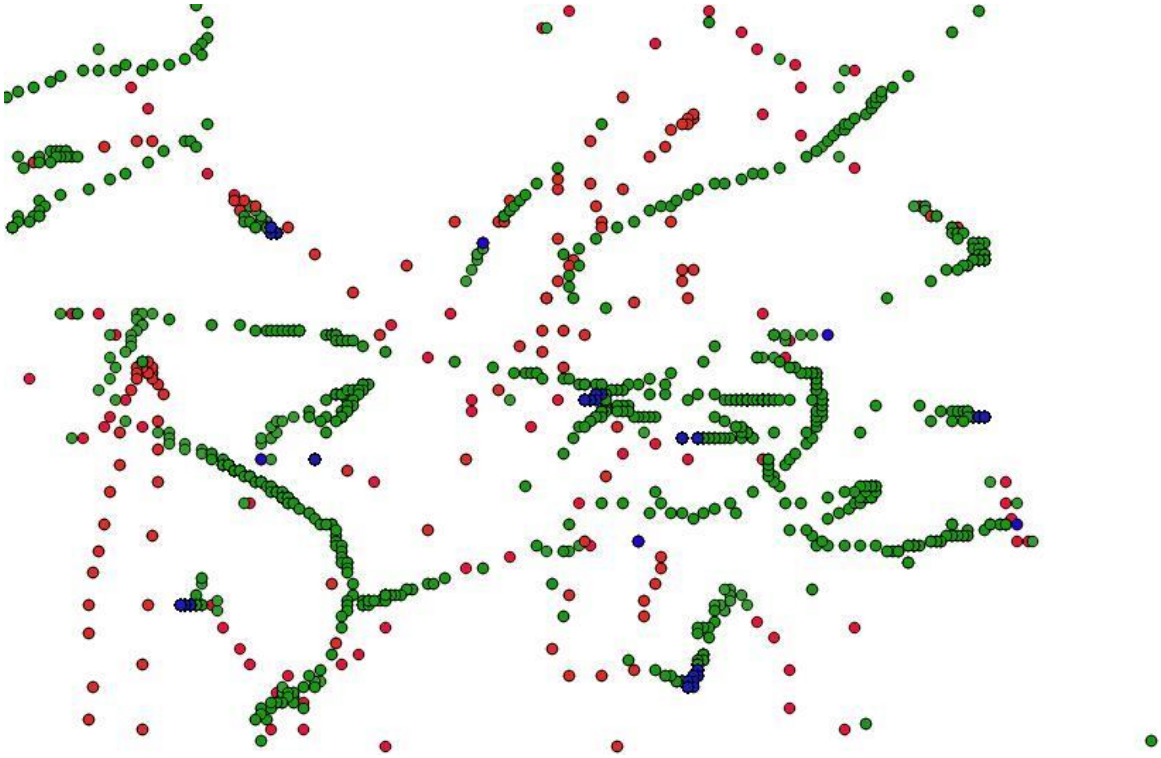


Figure 3.11 Raw behavioural inference of 1-2 s interval data from 10 collared animals at Penglaneinon. Behavioural classes are: Moving (Red), Grazing (Green) and Resting (Blue).

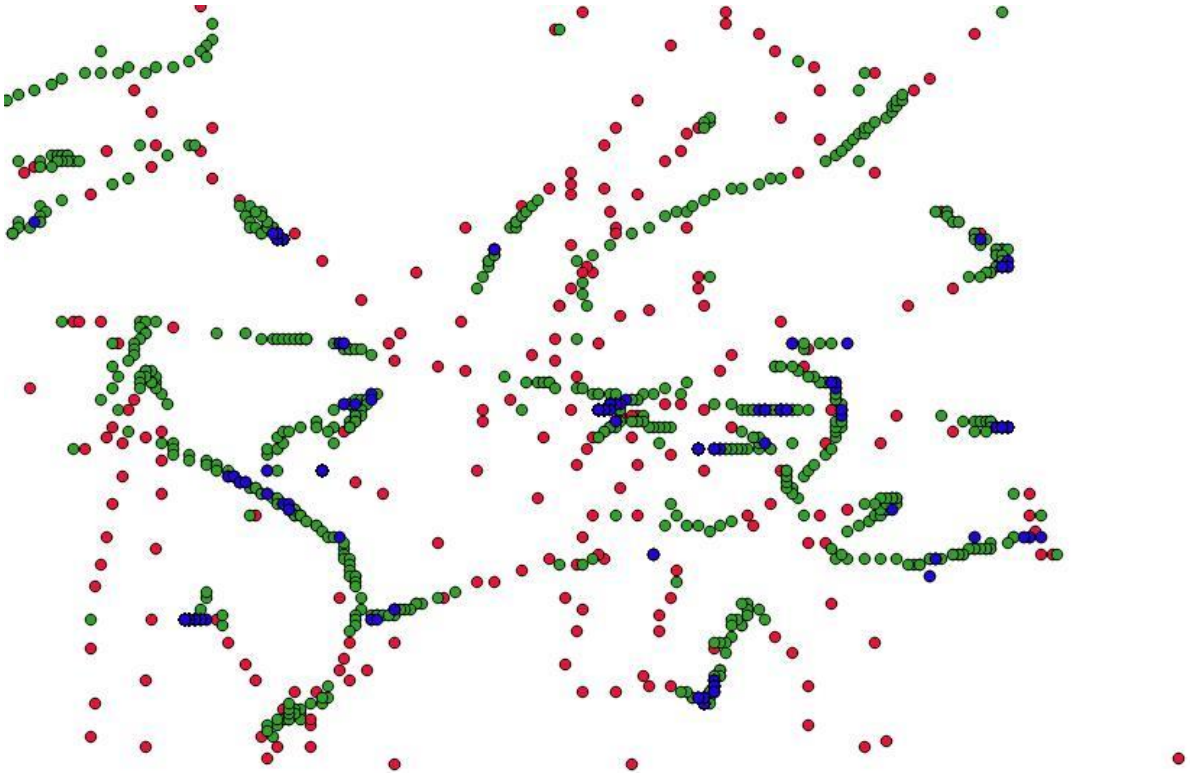


Figure 3.12 Averaged (30 s) behavioural inference using 1-2 s interval data from 10 collared animals at Penglaneinon. Behavioural classes are: Moving (Red), Grazing (Green) and Resting (Blue).

3.3.5 Section summary

Logger cost and battery life have long been considerable constraints in animal tracking operations. The incorporation of scheduling capability into simple low cost loggers design, as demonstrated in this study, can mitigate these constraints. Understanding the non-academic nature of many conservation-related tracking situations, the development of the Beferrer application provides a compromise between the desire for accurate focused behavioural analysis and accessibility to non-programmers. The level of customisation available in the application helps alleviate GNSS error and other erroneous recordings typical in these experimental low-cost loggers, whilst providing flexibility in the outputs available, and being widely applicable to tracking different animals in varied scenarios. The addition of the Schedsplitter application also allows a simple, quick and effective way of separating the recorded data into desired frequency intervals, which can then be used in further analysis.

3.4 Measuring grazing intensity across a site, and informing landscape decisions

3.4.1 Informed management decisions

The section will focus on addressing the fourth chapter objective *('testing the uniformity of grazing intensity across a given site'*), as well as provide examples to land practitioners of simple visual analyses which when combined with the behavioural classifications can help inform landscape based decision making.

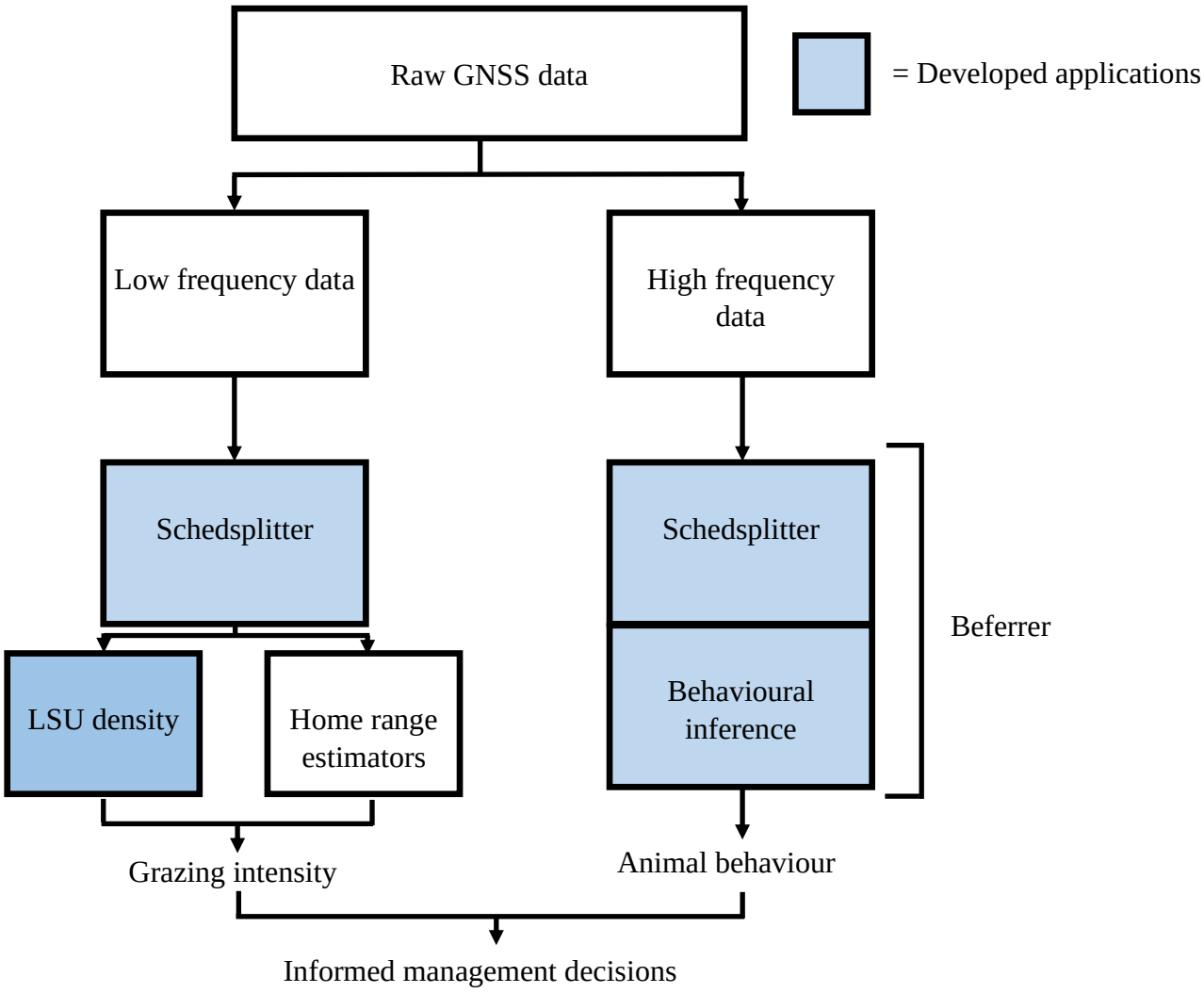


Figure 3.13 Flow chart of our proposed data handling and analyses steps.

The primary motivation for completing this objective was to investigate the effectiveness of LSU prescriptions which are used in many agri-environment schemes in the UK. One problem with the LSU approach is that it assumes uniformity in grazing intensity across a site. Were this not to be the case, the successful recovery of habitats would likely be stifled (as reported for the study sites used in this project). If sections of a given site are subject to over-grazing whereas others are being under-

grazed (or sometimes not grazed at all) then localised grazing pressure will deviate substantially from the prescribed levels. In order to adequately test this theory, existing analytical options were explored to see which could inform on site grazing intensity (i.e. home range analysis), whilst solutions were also developed to visualise intensity in relation to LSU values.

Home range estimators are widely used tools within the literature and have been proved successful in measuring the utilisation distributions (UD) of animals. There are also many existing software packages which make them readily accessible, even for non-programmers (e.g. Geospatial Modelling Environment (GME)). Two of the most commonly used tools to assess animal home ranges are kernel density estimators (KDE) and minimum convex polygons (MCP). KDE work by using a distribution function to calculate density of space use (Worton 1989) in which home-range boundaries are built by joining sites with equal density. In contrast, MCP determines the smallest polygon encompassing all external locations in a data set, with all interior angles less than 180° (Girard *et al.* 2002). KDE is one of the most statistically efficient methods for probability estimation, supported by a wealth of statistical literature, and the inclusion of many refinements and extensions (Fleming & Calabrese 2017). The use of MCP remains commonplace however (especially though the adapted local convex hull method), and related estimates have been used by many studies to draw conclusions about mammalian ranging behaviour (Nilsen, Pedersen & Linnell 2008). Its simplicity of use remains one of its major strengths (Burgman & Fox 2003).

However, as alluded to in the introduction to this chapter, the matter of which home range estimators should be used under different scenarios is matter for debate (Walter, Onorato & Fischer 2015; Péron 2019). The reasons for opting with these two methods in this study are as follows: 1) they are both simple, and easy to implement, and therefore suitable for land practitioners to use; 2) they are well recognised within the literature, and their effectiveness and flaws have been explored and documented in much detail; 3) while new third generation estimators such as auto-correlated kernel density analysis (AKDE) (Fleming *et al.* 2015) are capable of handling auto-correlated data), they are more difficult to implement. This leads to; 4) which is that rather than looking to infer specific behavioural or environmental interactions from these outputs alone, the aim is to simply compare differences in home range outputs from the two methods in how they infer grazing intensity across the two sites. Therefore, high precision in the home range estimation outputs is not required, and instead a more cautious use of them can be undertaken. This is line with Laver & Kelly (2008) who in their concluding remarks state '*managers using home range estimates in their decision-making should closely review the estimation methods without carte blanche acceptance thereof*''.

3.4.2 Home range estimator methods

Both KDE and MCP estimators were implemented using the 'adehabitatHR' package (Calenge & Fortmann-Roe 2014) in R Studio (R version 3.6.1) (Team 2013). Other packages used included 'dplyr' (Wickham *et al.* 2015), 'sp' (Pebesma *et al.* 2012), 'raster'(Hijmans & van Etten 2016), 'rgeos' (Bivand & Rundel 2017) and 'data.table' (Dowle *et al.* 2017). This code handled data processing and cleaning from raw data through to exported shape files from the home range estimators. The Schedsplitter.jar application was also called on during the analysis to separate the data into the desired intervals for processing. For the Ffridd Fawr site, only GNSS data from the first tracking session was included, as all the animals were swapped with a new flock at the end of the session (as per the farm's management regime) and therefore there could be no continuity in the animals measured from the first session. At Penglaneinon, data from both tracking sessions was included for the individual animals. Two-hour interval data was used for the analyses, with these selected for two reasons. Firstly, in their study using cattle, Pertotto-Baldivieso *et al* (2012) found that auto-correlation between successive observations could be minimised if time intervals were separated by at least 2 h. Secondly, given the obvious range restrictions of the sites in this study (by being completely enclosed), it was reasoned that the animals at both sites could reasonably cover the whole range available to them within an hour, and therefore the 2 h interval length allowed for full site coverage between recordings. Given the issue of frozen coordinates within the hourly interval data, these 2 h intervals were extracted from the schedule's eighth 'intensive' recording day as this data was intact. Therefore, a total of 12 recordings per 8 d was available for each animal. This resulted in a total of 359 observations for 7 sheep at Ffridd Fawr across a 49 d period (29/05/18- 16/07/18), and 542 observations for 5 sheep at Penglaneinon across a 107 d period (29/05/18- 10/09/18). Home range estimations for both merged animal data, and on an individual animal basis were completed. The reference/default bandwidth (H_{REF}) was selected for the KDE. Though 95%, and 50% isopleths are commonly selected to represent core and periphery areas of a home range, 10 percentile (i.e. 90, 80, 70, 60, 50, 40, 30, 20, 10) isopleths were considered most suitable for both KDE and MCP calculations in this study, as it would better allow grazing intensity to be viewed.

3.4.3 Home range estimator results and discussion

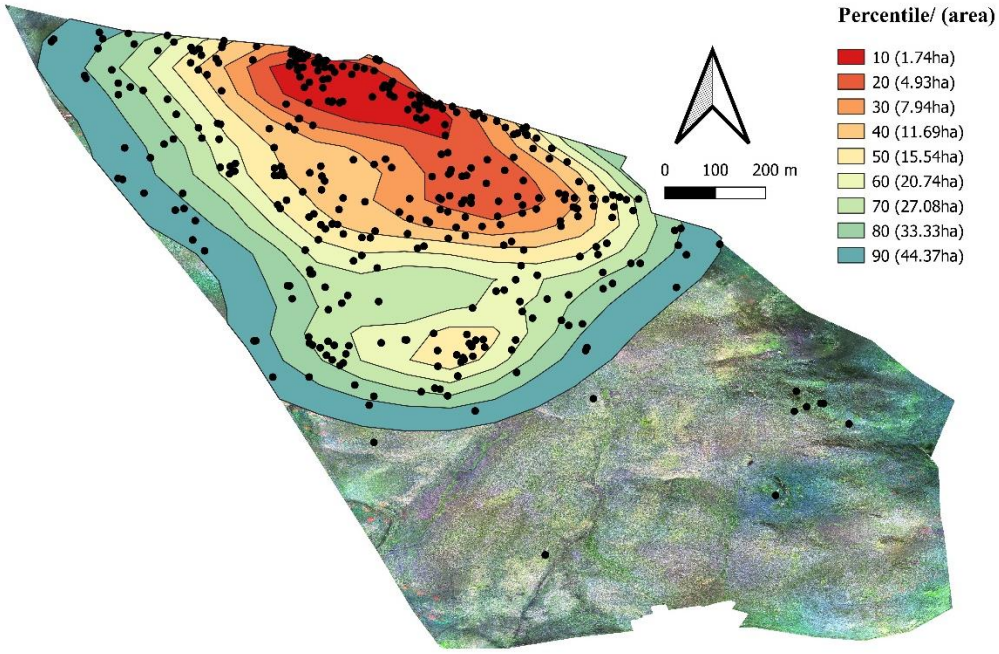


Figure 3.14 Combined Kernel density analysis at Ffridd Fawr site using 363 observations of 2 h interval data from 7 collared sheep. Isopleths are denoted in the legend in percentiles of 10, as well as the area (in ha) included for each isopleth. Base layer = narrowband RGB orthomosaic from UAV imagery.

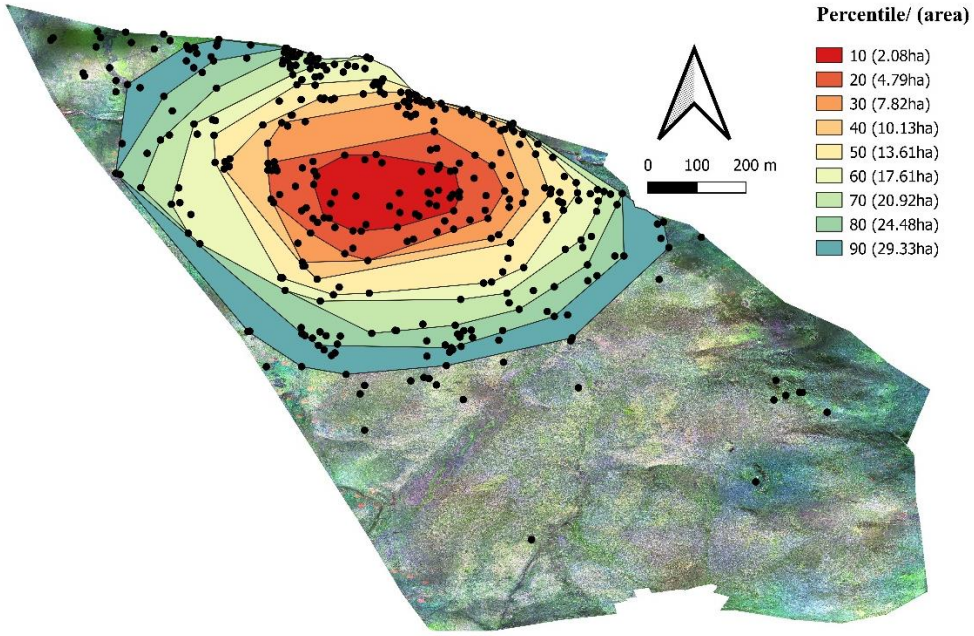


Figure 3.15 Combined Minimum convex polygon analysis at Ffridd Fawr site using 363 observations of 2 h interval data from 7 collared sheep. Isopleths are denoted in the legend in percentiles of 10, as well as the area (in ha) included for each isopleth. Base layer = narrowband RGB orthomosaic from UAV imagery.

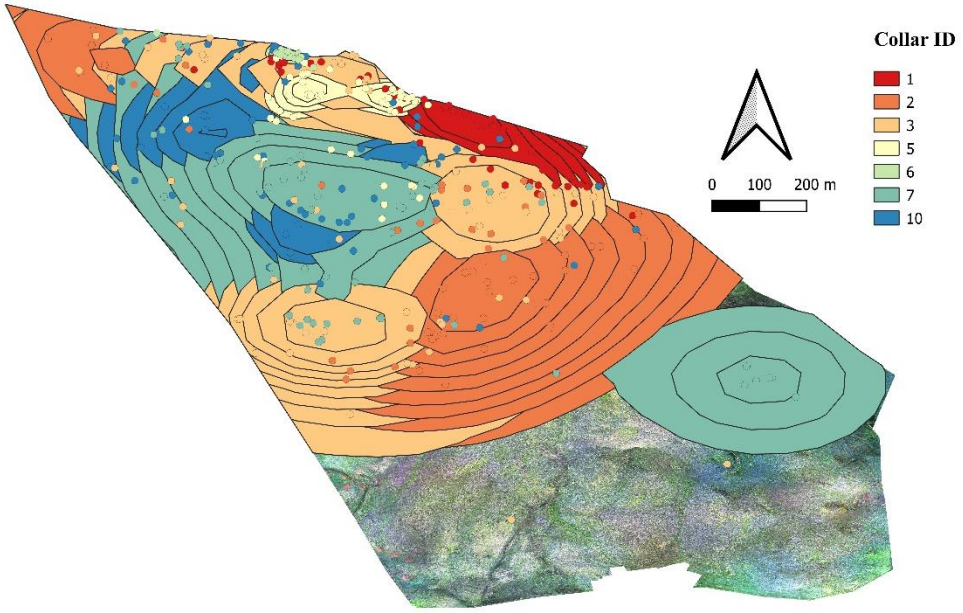


Figure 3.16 Individual Kernel density analysis at Ffridd Fawr site using 2 h interval data from 7 collared sheep. GNSS points matching the same colour as their individual home ranges are overlaid. Base layer = narrowband RGB orthomosaic from UAV imagery.

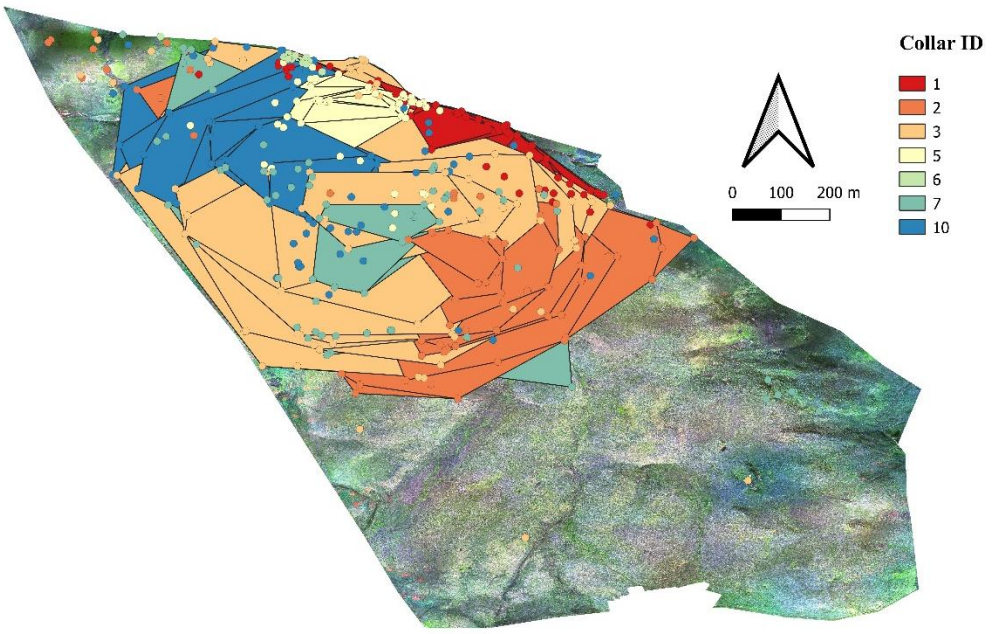


Figure 3.17 Individual minimum convex analysis at Ffridd Fawr site using 2 h interval data from 7 collared sheep. GNSS points matching the same colour as their individual home ranges are overlaid. Base layer = narrowband RGB orthomosaic from UAV imagery.

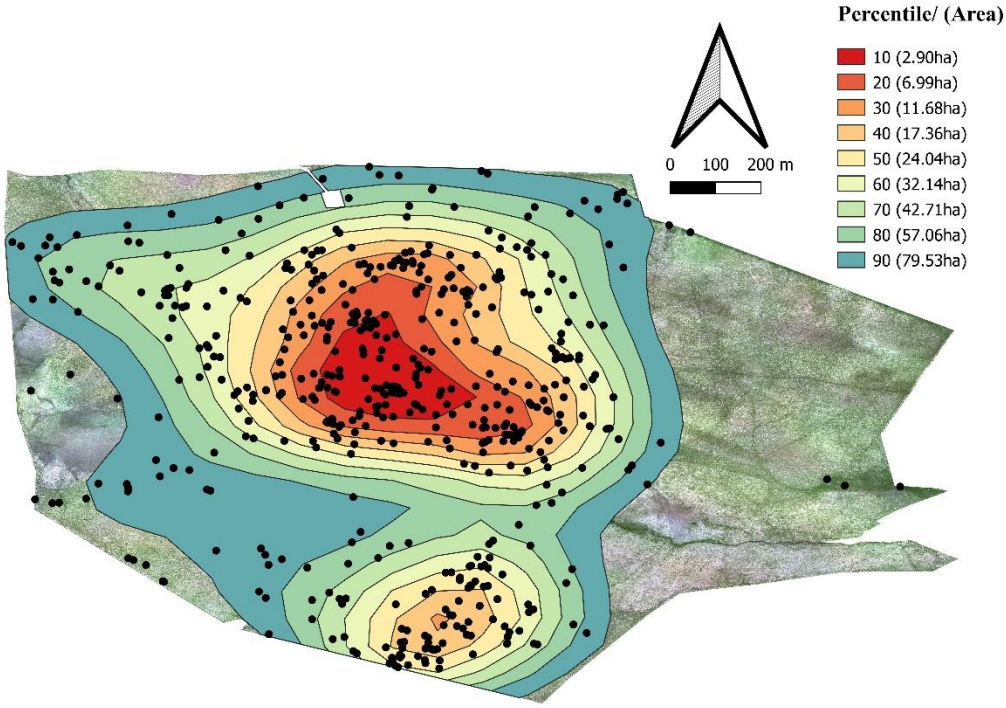


Figure 3.18 Combined Kernel density analysis at Penglaneinon site using 542 observations of 2 h interval data from 5 collared sheep. Isopleths are denoted in the legend in percentiles of 10, as well as the area (in ha) included for each isopleth. Base layer = narrowband RGB orthomosaic from UAV imagery.

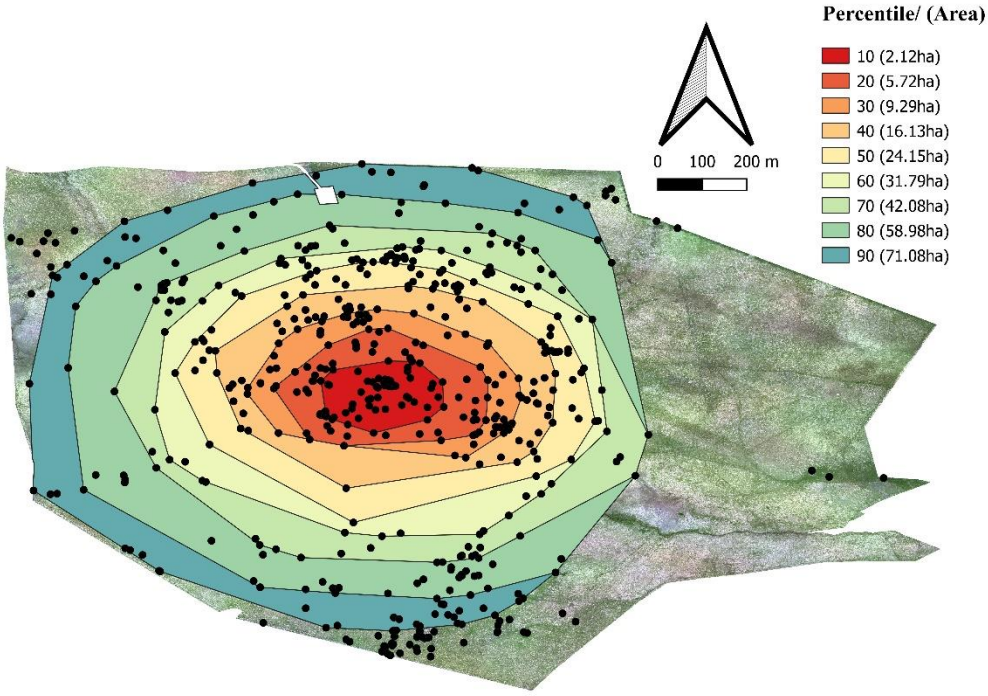


Figure 3.19 Combined minimum convex polygon analysis at Penglaneinon site using 542 observations of 2 h interval data from 5 collared sheep. Isopleths are denoted in the legend in percentiles of 10, as well as the area (in ha) included for each isopleth. Base layer = narrowband RGB orthomosaic from UAV imagery.

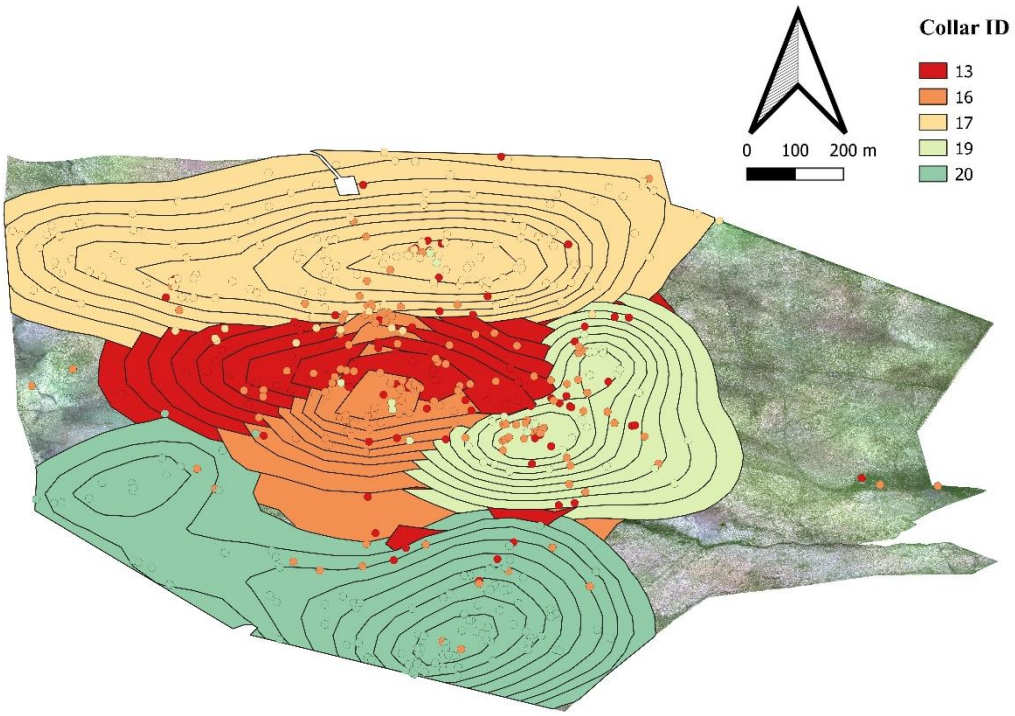


Figure 3.20 Individual Kernel density analysis at Penglaneinon site using 2 h interval data from 5 collared sheep. GNSS points matching the same colour as their individual home ranges are overlaid. Base layer = narrowband RGB orthomosaic from UAV imagery

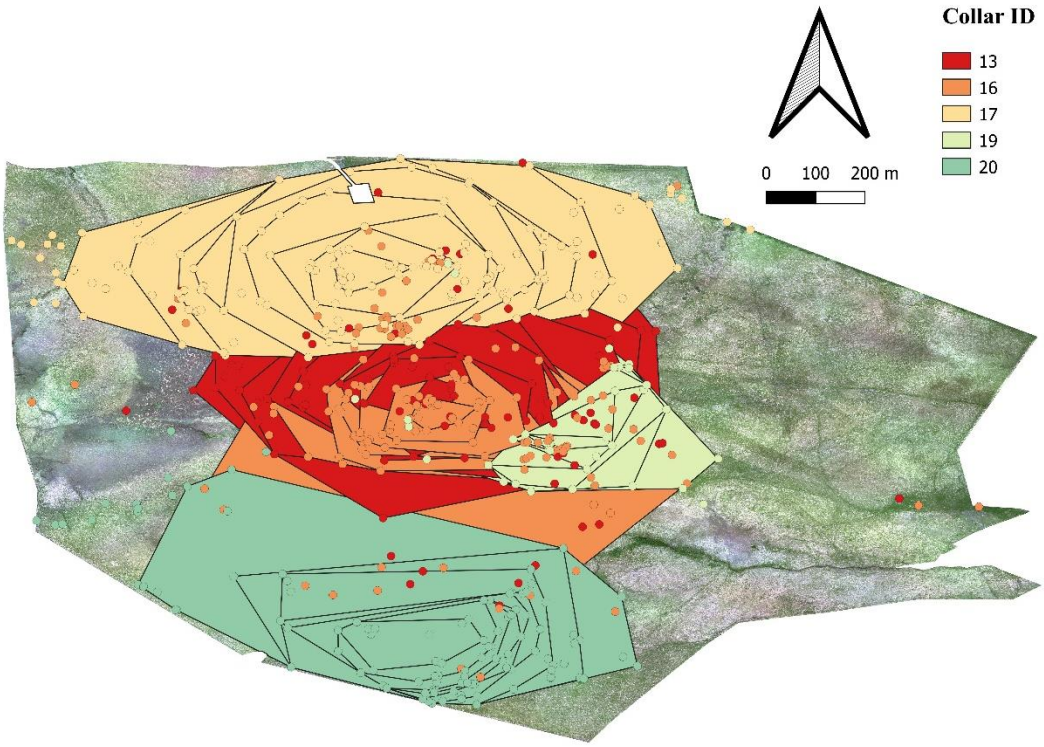


Figure 3.21 Individual minimum convex analysis at Penglaneinon site using 2 h interval data from 5 collared sheep. GNSS points matching the same colour as their individual home ranges are overlaid. Base layer = narrowband RGB orthomosaic from UAV imagery

The outputs from both the MCP and KDE indicate that grazing intensity across both sites did not include the whole utilisable available areas. At Ffridd Fawr, both the MCP (Figure 3.15) and KDE (Figure 3.14) outputs indicate the combined home range for all the animals was absent for the top half (southern end) of the site. They differ slightly however in that the KDE indicates the presence of a second smaller home range cluster above a larger one below it, whereas the MCP only indicates one. In the individual home ranges at Ffridd Fawr, there appears to have been considerable overlap between home ranges. However, there is a noted difference between the individual MCP (Figure 3.17) and KDE (Figure 3.16) results for Ffridd Fawr in that the KDE suggests that for one sheep (collar 7) its home range extended into the top half of the site (northern side), where the MCP does not. At the Penglaneinon site, the combined KDE (Figure 3.18) and MCP (Figure 3.19) outputs differ in the respect that the KDE indicates the presence of two home ranges clusters, whereas the MCP suggests a single home range cluster to operate in the centre of the site. With regards to the individual home range estimates at Penglaneinon, unlike Ffridd Fawr there appears to be less overlap between the home ranges. Both the MCP (Figure 3.21) and KDE (Figure 3.20) indicate that two sheep in particular (collar 17 and 20) are likely occupying opposite ends of the site (north and south), whilst the remaining three sheep occupy the central area. This could suggest the presence of three potential home range clusters at least at Penglaneinon, though more data and individuals collared would be required to confirm this. Finally, as we can see from Table 3.4, in all but three of 10 percentile isopleths, the KDE area was greater than the MCP estimates, though the difference at the lower percentiles was minimum, which suggests a strong level of consistency in the outputs. However, towards the higher percentile isopleths (80, 90, 95), the degree of difference became far more apparent, which would suggest a lower level of certainty in the home range estimation at these increments. That said, at both sites it appears the inclusion of a second home range cluster in the KDE is likely to have contributed at least partially to this difference.

Table 3.4 Comparison of kernel density estimation (KDE) and minimum convex polygons (MCP) analyses outputs from both study sites. Blue values in last column, where KDE outputs are greater than MCP outputs. Gold values in last column, where MCP outputs are greater than KDE outputs.

Site	Percentile (%)	KDE area (ha)	MCP area (ha)	KDE – MCP (ha)
Ffridd Fawr	10	1.737859	2.08359319	-0.34573374
	20	4.929966	4.79048618	0.139480127
	30	7.940678	7.82197524	0.118702905
	40	11.69265	10.127832	1.564813257
	50	15.54438	13.6072756	1.937107324
	60	20.74205	17.6063521	3.13570127
	70	27.07908	20.9157459	6.163338892

	80	33.33131	24.4807427	8.85057041
	90	44.36783	29.3295176	15.03831211
	95	55.38285	31.94779	23.43506
Penglaneinon	10	2.898938	2.12165427	0.777283789
	20	6.992676	5.72305857	1.26961748
	30	11.68018	9.28966409	2.390512891
	40	17.35795	16.1253964	1.232557642
	50	24.0425	24.1454208	-0.10291638
	60	32.14288	31.7905781	0.352304053
	70	42.71297	42.0845732	0.628394897
	80	57.05567	58.9814329	-1.92576013
	90	79.53098	71.083923	8.44705498
	95	96.6464	77.12527	19.52117927

The results for home range analysis provide clear indication that the grazing pressure at both sites is not uniform. However, the isopleth nature of the home range analysis outputs does not provide suggestions as to how grazing intensity compares locally across a site. This is important, because to best inform on management decisions, we must be able to view grazing intensity at a local scale (which is difficult to do using isopleths). Also, in order to assess whether current management prescriptions are effective, there must be a comparison to the metric of which they are being applied - which in the case of the sites in this study and many others in the UK, is LSUs.

3.4.4 Livestock unit estimator method

A simple method of estimating true LSU coverage across a site was devised according to the positional locations recorded. Given the relevance of this output for land practitioners an easy, non-programming based approach using the open-source GIS software QGIS (vers 3.4.15-'Madeira') (QGIS development team, Team 2016) was opted for. Firstly, a grid (1ha cell size (for LSU comparison)) was constructed (Vector - > Research tools - > Create Grid). This grid was then clipped to the extent of the site using the site 'mask' (Vector - > Geoprocessing tools - > Clip). Area (in ha) was then calculated for a new column in the attribute table (layer properties - > Open Attribute Table - > Open Field Calculator (Expression-$Area/10000)). Next the 2 h interval GNSS data was loaded into QGIS (Layer - > Add Layer - > Add Delimited Text Layer), and the number of observations in each grid counted (Vector - > Analysis Tools - > Count Points in Polygon). Finally, to calculate LSU's based on point density in the area of each grid, the below formula was created and applied into

a new column of the attribute table (layer properties - > Open Attribute Table - > Open Field Calculator (Expression-*Equation*);

$$((N_A/N_o)*N_p)*LSU_A)/A$$

In this case, N_A is the number of animals on a site (i.e. 180 in the case of both sites), whilst N_o is the number of recorded observations used (i.e. 542 for Penglaneinon, and 359 for Ffridd Fawr). N_p is the number of observations calculated for each polygon (usually labelled 'NUMPOINTS' in the attribute table), whilst LSU_A is the livestock unit value for an animal type used (i.e. sheep in our case = 0.15 (using unit values from the agri-environment scheme '*Glastir*')). Finally **A** is the area (in ha) calculated within the attribute column for each polygon grid cell.

What this output provides is the cumulative LSU estimate per cell of all animals for the time period of the observations recorded (i.e. 107 days at Penglaneinon, and 49 days at Ffridd Fawr). Like LSU values, the values we use are calculated per hectare basis. Therefore, although all cells not on the edge of the site are a hectare in size, the size of cells along the site edge obviously vary and therefore the values within these are scaled to a per hectare estimation.

3.4.5 Livestock unit estimator results and discussion

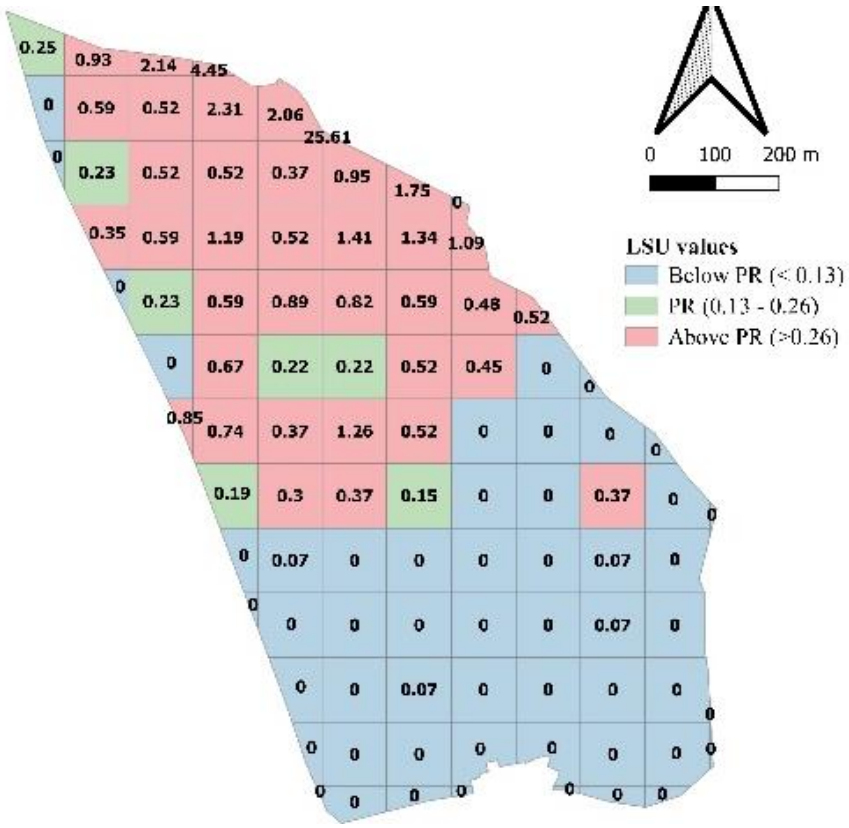


Figure 3.22 LSU estimation at Ffridd Fawr site based on point in polygon counting using combined GNSS observations from all animals. Internal grid cells = 1 ha. PR denotes the 'prescribed range' in LSUs for the site under 'Glastir' management.

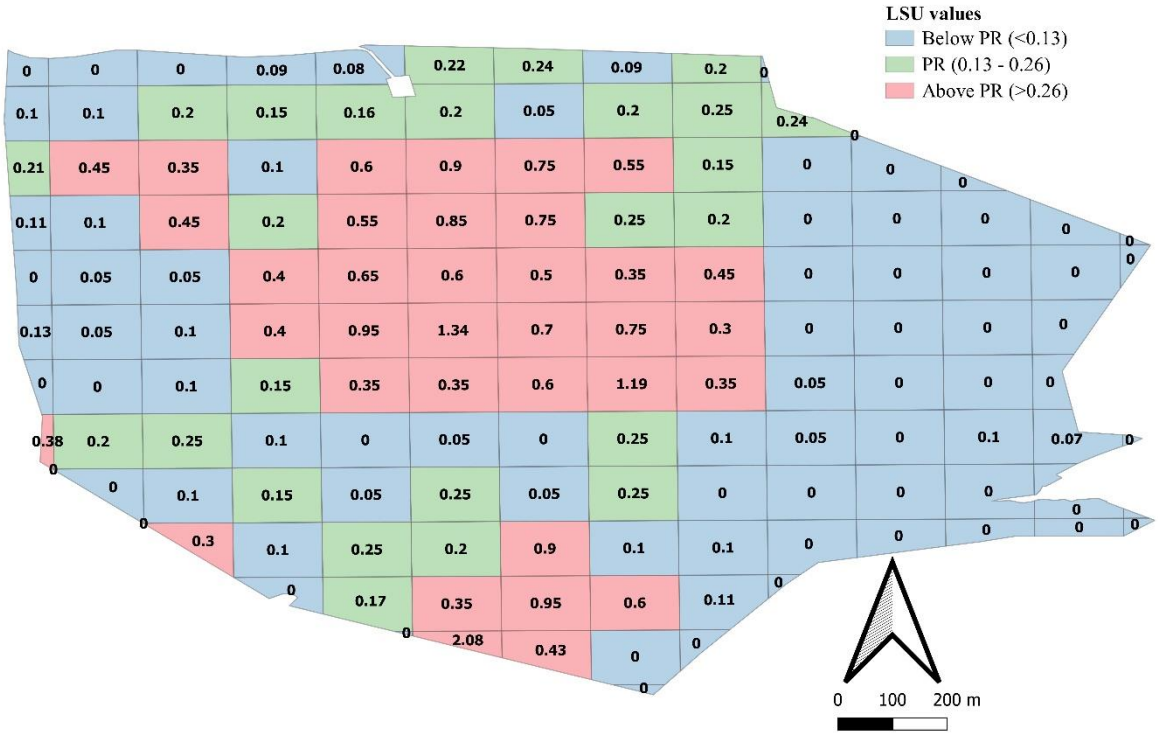


Figure 3.23 LSU estimation at Penglaneinon site based on point in polygon counting using combined GNSS observations from all animals. Internal grid cells = 1 ha. PR denotes the 'prescribed range' in LSUs for the site under 'Glastir' management.

Table 3.5 Comparison of LSU estimations using GNSS data at both sites.

	Ffridd Fawr		**Penglaneinon**	
LSU range	**Number of cells**	**Percentage (%)**	**Number of cells**	**Percentage (%)**
No presence	51	51.5	52	36.9
Below prescribed range	55	55.6	80	56.7
Prescribed range	7	7.1	26	18.4
Above prescribed range	37	37.4	35	24.8
Total number of cells	99		141	

The outputs of the LSU estimation concurred with those of the home range analysis in that grazing did not appear to include the whole site or be uniform in manner. At both sites (Figure 3.22 and Figure 3.23) the percentage of cells within the prescribed LSU range was low (18.4% and 7.1%).

Furthermore, the general indication at both sites is that they were generally under-grazed, given the percentage of cells with LSU values below the prescribed range was greater than the values above (56.7%/55.6% against 24.8%/37.4%). The main difference between the sites is that for Ffridd Fawr, 92.7% of the cells (51/55) below the prescribed range contained no presence, whereas at Penglaneinon only 65% of the cells below the range had no presence.

When viewed with the home range outputs (Figure 3.24 and Figure 3.25) a clear overlap between the two types of analyses is evident, and provides confidence in the interpretation. However, currently the points-in-polygon calculation for the LSU estimate does not account for the number of different individuals within a cell but only the density of points. Future work would therefore need to focus on deriving an alternative counting method which included a weighted approach to account for both number of individuals and their respective point density. Furthermore, the current method assumes a continuous grazing intensity over the entire tracking period, whereas in reality it could be that in some cases there is an intensive occupation period for a small amount of time (e.g. 24 hr.), whereas another cell may contain the same amount observations within a cell but these recordings span the whole observation period. Given the difference in effect short intensive grazing periods have on vegetation compared to longer lower intensity grazing (Johnson 1953; Adler, Raff & Lauenroth 2001) - accounting for the temporal intensity would also be a key improvement going forwards.

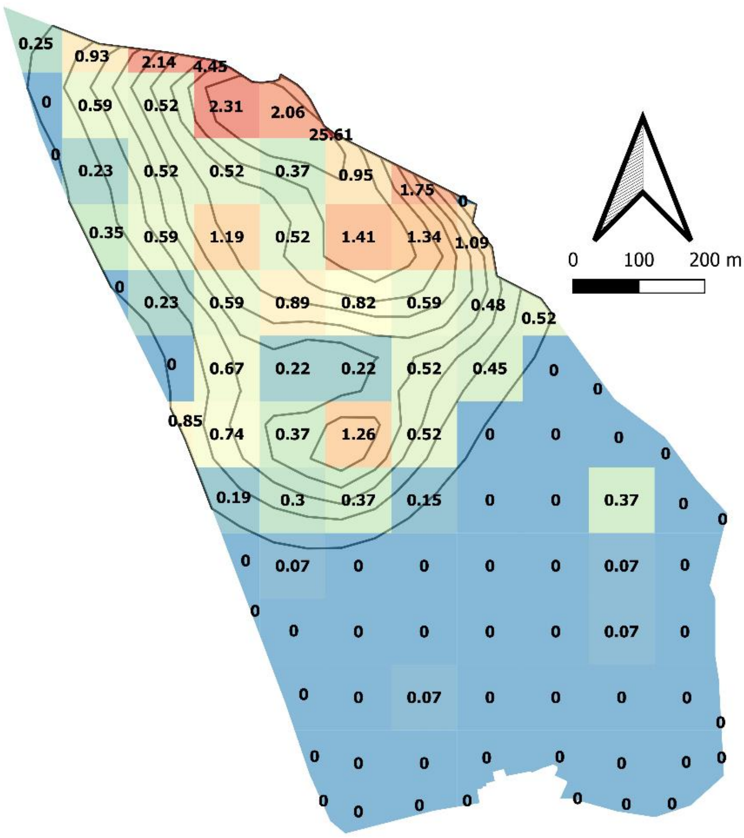


Figure 3.24 LSU estimation at the Ffridd Fawr site with combined KDE contour lines representing 10 percentile isopleths. LSU estimation was completed using point in polygon counting using combined GNSS observations from all animals. Internal grid cells = 1 ha.

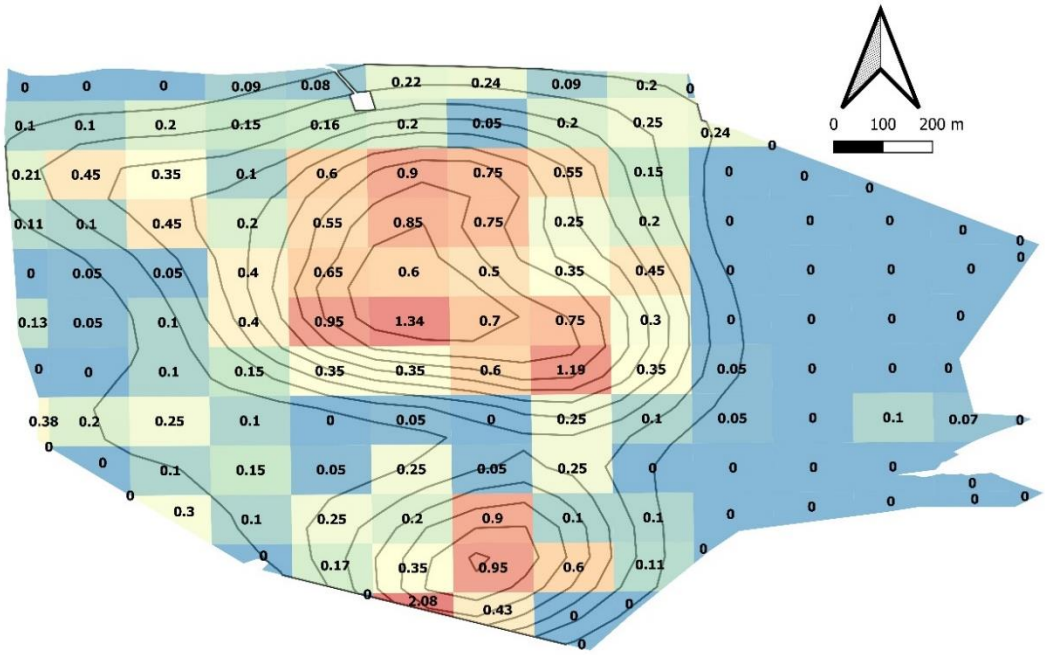


Figure 3.25 LSU estimation at Penglaneinon site with combined KDE contour lines representing 10 percentile isopleths. LSU estimation was completed using point in polygon counting using combined GNSS observations from all animals. Internal grid cells = 1 ha.

3.4.6 Informing management decisions from combined information layers

Though it is clear that further validation and improvement to the LSU estimation method is required before there can be surety in the outputs produced, when combined with the home range outputs, they do provide an effective tool for assessing site distribution patterns. Furthermore, when the outputs are viewed alongside the behavioural analysis data (such as environmental layers and other information sources) - possible explanations as to the distributions witnessed may be investigated.

Figure 3.27 and Figure 3.26 are but two examples of how combining information layers together in GIS software can provide meaningful observations as to how a given site is being utilised, and possible indications as to why.

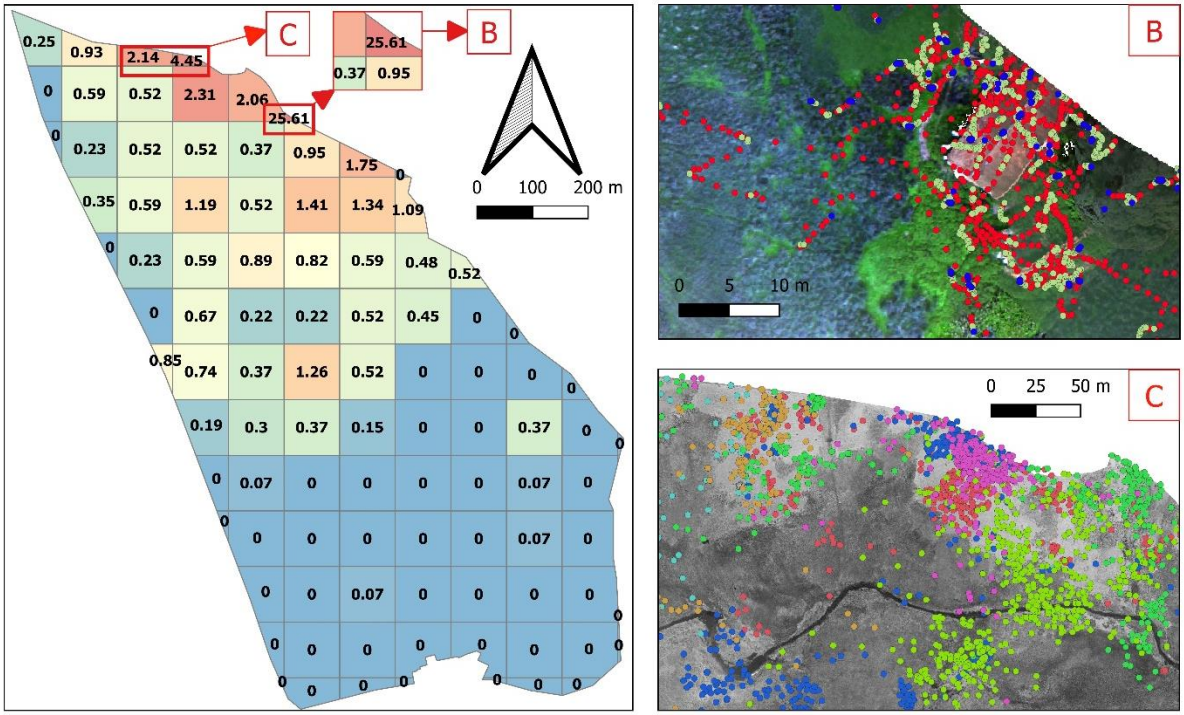


Figure 3.26 Combined information layers for Ffridd Fawr. Main image (left) = LSU estimation based on 2 h GNSS data. B = 1 s interval behavioural classes from GNSS data overlaid on top of an RGB presented UAV derived multispectral image. Within B; Green Dots = foraging; Red dots = moving; Blue dots = resting. Behavioural classes were generated using Beferrer. C = 5 min interval GNSS data overlaid on top of an NDVI layer.

At Ffridd Fawr (Figure 3.26), in the center of the KDE output, a number of cells with high LSU values (one in particular at 25.61 LSU) are visible. When additional layers are utilised, it is evident from image B that in the center of this unusually high LSU cell is a partially derelict building. From the 1 s behavioural data a clear high level of activity in and around the building is observed. This would strongly suggest that the animals are using the building as shelter, and that this was a key reason as to why distribution was so high in this area. Looking at the other highly occupied cells, many of the observations within image C are in the patches where there are higher NDVI values. This would further indicate that the animals were drawn to the greener semi-improved pastures available in these areas.

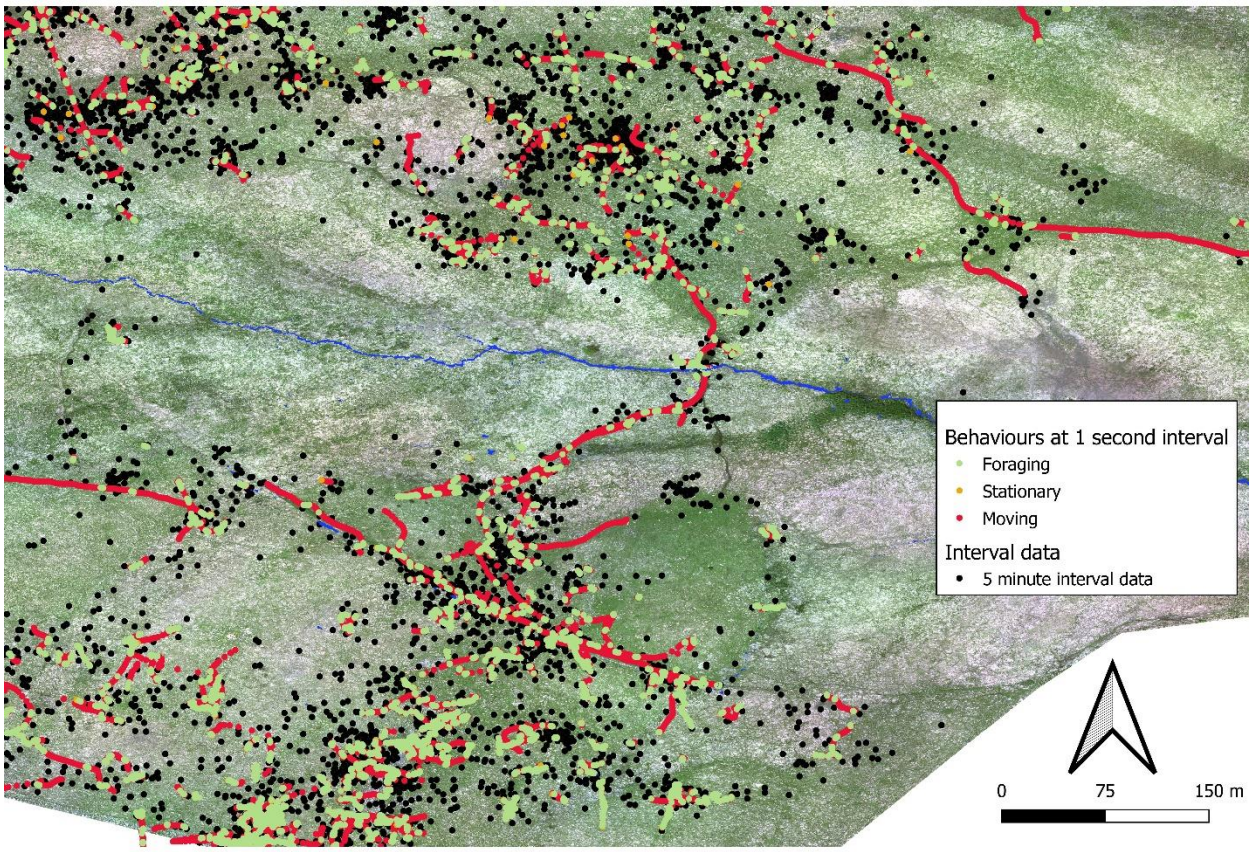


Figure 3.27 Behavioural classes from 5 collared sheep. Blue line = running stream exemplified by a standing water mask.

The LSU and home range maps for Penglaneinon indicate two concentrate zones where animals have been spending their time. By combining 1 s behavioral analysis results from Beferrer, as well as the 5 min interval data and two of the environmental layers, an indication is provided as to why the distribution may be so. Looking at Figure 3.27, the high frequency movement data supports the presence of these two clear activity zones. Beyond that, much of the observations in the top zone are located clearly within the area that was mown two years ago, and that animals are moving between the two zones using a single crossing point over the stream.

Through combining different interval frequency data with the simple analysis tools presented, alongside other environmental layers, it is possible to infer much about the animal distribution on given a site through visual interpretation only. This approach offers much to land practitioners, and provides a simple, but effective means of informing land management decisions. However, given the lack of statistical rigor, it would be ill advised to overstate the results of these analyses. Instead these methods are best utilised as broad descriptors of site utilisation, and indicators of potential resource importance. There therefore remains a need to include a robust statistical approach to quantify the animal-environment interactions occurring on a given site, which will be addressed in the following chapter.

3.4.7 Section summary

This section focused on addressing the fifth objective *('testing the supposed uniformity of grazing intensity across a given site'*) and provided examples of how simple visual analyses could be used to help inform landscape-based decision making. Both the home range analyses, and LSU estimator concurred that grazing pressure at both sites was uneven, with most areas indicated at being under-grazed (55.6% and 56.7% respectively). Using these outputs alongside the developed 'Beferrer' behavioural inference software and other environmental layers, a considerable number of facets surrounding the sheep behaviour on each site could be deduced through simple visual interpretation of differing data combinations. Though more robust statistical approaches (explored in Chapter 4) are required to get a complete picture, the simple analytical techniques presented in this section are effective tools in exploring the characteristics of land usage by animals on a given site.

3.5 Chapter conclusions and areas for future research

When attempting to track large numbers of animals it is crucial to have in mind the biological objectives that need delivering. It also important to consider factors such as the length of the tracking session (and therefore battery life required), the frequency at which data will need to be collected, and at what resolution. Finally, the cost of tracking unit must be considered since this may dictate whether enough animals have been tracked for the data to be representative.

From a hardware and firmware perspective, the approaches developed in this study provide a low-cost solution to tracking large number of animals effectively and efficiently. Unfortunately, time constraints and the limitation of available resources meant that large-scale observations of all animals on both sites could not be completed – however, the validation of the UAVRTS has proven it to be capable of performing such a task, and at a low cost. Used correctly this would be an effective way of collecting site utilisation data for analyses such as home range estimators or the LSU estimator. The incorporation of a moderate number of GNSS loggers could then be used to fill the temporal voids outside of the radio tracking system, as well as inform on specific individual behavioural patterns. This type of integrated tracking system approach could provide new insights into movement ecology. The ability to flexibly alter the different hardware utilised for any given study will result in greater reliability in the data gleaned, and improve the efficiency of the tracking process. Other tracking approaches which could complement these methods are already being explored. In one example, an undergraduate project explored the potential of image recognition software to identify animals from the UAV imagery derived as part of the current project (Yearby 2018). This has the benefit of not requiring any equipment to be used on the animals, but could also be used in conjunction with the UAVRTS to improve the accuracy of the observations, whilst using the identification capability of the UAVRTS to identify specific animals.

Finally, the data processing software (Beferrer and Shedsplitter) developed allows land-based practitioners to access data insights which have in the past been predominately restricted to academic studies. Though further work could be done to improve the usability of the software (e.g. inclusion of a graphical user interface (GUI)), these tools are already being used in new tracking work at Ffridd Fawr under the RSPB, whilst follow-on academic based tracking projects at the university have also benefitted from them.

CHAPTER 4

SPECIES DISTRIBUTION MODELLING

4.1 An introduction to ecological niche theory and species distribution modelling

4.1.1 Understanding the role of ecological niche theory in species-environment analyses

Analysis of species-environment relationships is a key issue set firmly at the heart of predictive geographical modelling (Guisan & Zimmermann 2000), and is founded on the principles laid out in ecological niche theory (Grinnell 1917; Elton 1927; Hutchinson 1957). Ecological niche theory conveys the relationship a species population has to all aspects of its environment (Khatibi & Sheikholeslami 2016), which broadly separates into two related components: (1) the biotic and abiotic factors that influence a species given locality; and (2) the impact of species on those factors (Leibold 1995).

The concept further describes two niche types: the fundamental niche ($N_{\mathbf{f}}$) and the realised niche ($N_{\mathbf{r}}$). The fundamental niche comprises the area an organism would have if there were no restricting factors (e.g. competitors/predators, physical boundaries, disease), whilst the realised niche refers to the actual environmental space that an organism inhabits, and the resources it can access as a result of limiting pressures (Franklin 2010a; Sillero 2011).

4.1.2 The difference between mechanistic and correlative approaches

Mechanistic models are used to parametrise the fundamental niche, whereas correlative statistics exist to assess the realised niche. Of the two approaches, the most widely used are correlative models as they more simply seek to statistically link spatial data (environmental variables) to species distribution records, whereas mechanistic approaches rely on detailed physiological sampling techniques to investigate the functional traits of an organism to its environment (Kearney & Porter 2009). Regardless of resource practicality, the decision as to which approach is utilised (either mechanistic or correlative) should ultimately be decided by the specific objectives and circumstances of the study undertaken. Within the context of this study, the prime objective has been to estimate the occupied area (realised niche) of each site, then rank the effect of the selected environmental variables (both direct and indirect) on the distribution of the animals, so manipulation can then take place. Though a mechanistic understanding of the underlying physiological responses of the animals to the environment would be insightful, it is not required to address the research questions at hand, and as such a correlative approach is more prudent.

4.1.3 Introducing species distribution modelling

Species distribution modelling (SDM) is the hypernym used for the broad groups of statistical methods used to describe empirical correlations between species distributions and environmental

variables (Franklin 2010a). Other terms associated include 'ecological niche models (ENM)' (Harrison 1997), and 'biological envelope models' (Araújo & Peterson 2012) which are used interchangeably in many cases in the literature. The use of these terms still engenders debate regarding the exact varying individual nuances that separate them (Sillero 2011; Peterson & Soberón 2012). The applications for SDMs are widespread with examples including (but not limited to); assessing species invasion and proliferation, conservation planning and reserve management, quantifying environmental niches of species, and assessing the impacts of climate, land use and other environmental changes on species distributions (see Guisan & Thuiller (2005) for further examples).

4.1.4 Static vs dynamic modelling

As surmised from the list of applications, the main enticement for users is the predictive nature of the models. This includes temporal (i.e. species distribution responses in the future, e.g. as a result of climate change) and spatial predictions (i.e. projecting the calculated niche of a species to a different geographic space, e.g. to see if an invasive species is likely to proliferate elsewhere) (Elith & Leathwick 2009). These predictive modelling techniques can be either static or dynamic in nature (Beerling, Huntley & Bailey 1995). Static models provide time-independent equilibrium predictions, whereas dynamic models predict time-dependent dynamic responses to a fluctuating environment (Robertson *et al.* 2003). Though dynamic models possess many advantages over static techniques (given they compensate for environmental change through time), they are more complex in both design and implementation with considerable input required to achieve reliable outputs (Guisan & Zimmermann 2000; Franklin 2010b). In the context of this study, dynamic modelling could potentially offer further insight than a static approach, however it is not essential to deliver the prime objective of the chapter (*'ranking the environmental variables affecting sheep distribution'*). Thuiller *et al* (2008) notes that; '*a trade-off exists between complexity and tractability in modelling species and identifying the most reliable and unbiased solution is not a trivial task'*. As such exploring the pros and cons of such an approach would be worthwhile in a future study, but it is not necessary or prudent in the current context.

4.1.5 The importance of the conceptual framework for accurate modelling

The use of a static correlative species distribution model represents the most appropriate method of delivering the prime objective of this study. As suggested previously, though predictive maps are often considered the main output of interest for most, these SDMs also allow variable importance for each predictor variable (both direct and indirect) to be calculated and ranked. As such, it is this output which will be vital to conclude the study. Despite being simpler in their approach than mechanistic-dynamic models, these correlative models are often executed improperly, with confusion even among the literature with regard to the published limitations of different data, and the uses that various outputs can justifiably support (Guillera-Arroita *et al.* 2015). It is therefore crucial to base the

formulation of an SDM on an underlying conceptual framework (Guisan & Zimmermann 2000). Though the exact number of steps often varies between authors, the first four stages are synonymous: 1) conceptualization, 2) data preparation, 3) model fitting, and 4) model evaluation (Guisan & Zimmermann 2000; Guisan & Thuiller 2005; Kearney & Porter 2009; Franklin 2010a). Further steps such as 5) 'spatial predictions' (Guisan & Zimmermann 2000) and 6) 'model applicability' (Guisan & Thuiller 2005) are often incorporated to assess the ability of a model to project into the future and in new areas. However, as such projection will not be undertaken in the current study, these further steps are not required in this context. Given the process is often iterative, where certain model steps may be revisited and improved, the process can be viewed cyclically rather than with a predefined endpoint (Zurrell 2019).

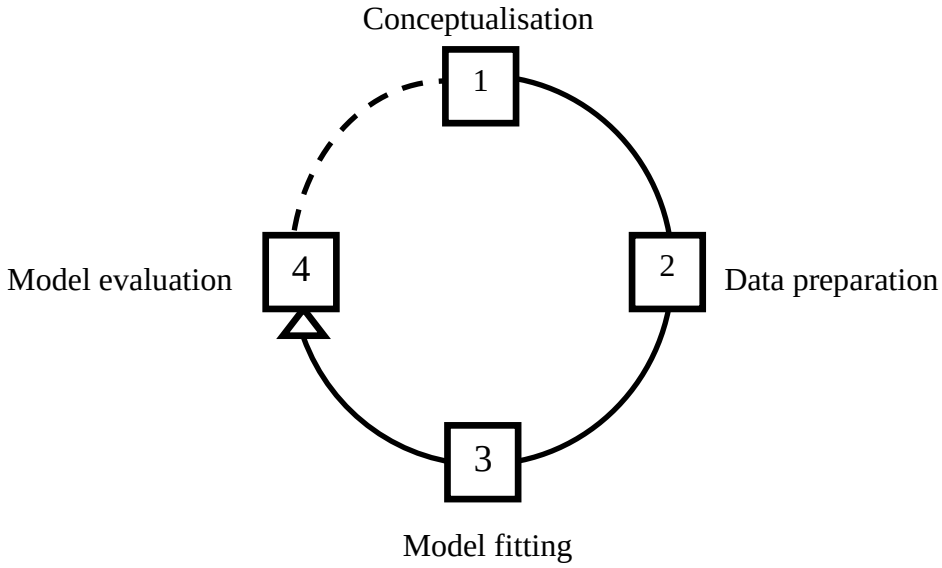


Figure 4.1 The species distribution model process viewed cyclically. Modified from Zurrell (2019).

4.2 A checklist for developing the models correctly

The following sections will provide detail on the required processing steps for the modelling to be undertaken correctly, and how these are to be accommodated within this study.

4.2.1 Conceptualisation

The conceptual model is fundamentally concerned with the ecological concepts that underpin the study (Guisan & Zimmermann 2000). Guisan & Thuiller (2005) suggest a series of points that require addressing for conceptualisation to have been fully considered. These are considered in the following points

4.2.1.1 The assumptions behind models and niche concepts

A fundamental assumption employed within the static SDMs is that there is equilibrium (or at least pseudo-equilibrium) between the environment and the observed species present (Guisan & Zimmermann 2000; Austin 2002). This implies that the study population present is not new to the study area, and/or not expanding its range or distribution (e.g. in the case of invasive species) (Robertson, Villet & Palmer 2004). In the context of this study, the presence of a physical boundary (fence) guarantees that the sheep are not actively expanding their range, and the fact that most sheep have been present on the site before (Chapter 2) indicates a settled state, which is further emphasised by the indication of hefting occurring (Chapter 3). The suggestion that challenges the equilibrium postulate within most studies is the notion that climatic or environmental change through time alters the behaviour of the observed species (Guisan & Thuiller 2005). By this reasoning, any inference of behavioural response to environmental variables can only represent a 'snapshot' in time (Guisan & Theurillat 2000), as subsequent changes in responses to the environment are likely to occur given the dynamic nature of most systems. This is particularly problematic for studies wishing to create accurate predictions on species distributions or their responses in the future. Conducting animal manipulations as a result of previous modelling would of course be problematic with regard to the assumption of equilibrium, as variable importance may well alter temporally. However, for the current study this is not the objective, though recommendations for manipulation with consideration to the equilibrium postulate will be outlined in the discussion.

Aside from the assumption of equilibrium, the possibility of biotic interactions (primarily through competition or predation) must be considered (Guisan & Thuiller 2005). This is because static SDMs do not usually account for species interactions (unless included as a predictor variable), and therefore over-or under-estimation of environmental responses can occur (Franklin 2010a). On both sites within this study, competition is likely to be minimal as only single species grazing management is occurring (e.g. there are no cattle or ponies grazing with the sheep). Furthermore as local populations are extremely scarce, encroachment by large herbivores (such as deer) is highly unlikely, and therefore any grazing competition will realistically only be from *Oryctolagus cuniculus* (European rabbit), or *Lepus europaeus* (European hare), and be minimal in comparison to the level of sheep grazing occurring. Likewise, the level of predation is also likely to be negligible as all study animals are mature individuals and no natural predators occur within the study areas. Interactions with humans would likely represent the greatest influence on distribution, but as animals were only handled twice within the study period with no supplementary feed given, the related impact was again likely to be small. However, at the beginning of the tracking period, as lambs are present with ewes this may cause a pseudo predatory awareness to manifest in their behaviour.

4.2.1.2 Assessing available and missing data, the relevance of environmental predictors for the focal species and the given scale, and deciding on the appropriate sampling strategy.

Available and missing data, the relevance of environmental predictors for the focal species and the given scale, and deciding on the appropriate sampling strategy were all considerations extensively explored and discussed within Chapter 2 of the study: 1) the justification of the UAV derived layers was clear once existing data, other sampling approaches, and the required spatial scale had been assessed; 2) the environmental variables represented in the assembled GIS layers were based on the specific ecological considerations for each site. The limitations of these environmental layers were also considered.

Within SDM terminology, Austin (2002) defined environmental predictors (factors) as either proximal (direct/causal) or distal (indirect/proxy) with regard to their effect on species distributions. Proximal factors correspond to predictors which directly affect the physiology of the observed species, either through regulation (e.g. temperature), or as resources (e.g. food, water). Distal factors on the other hand relate to resources or regulators, but only indirectly and through correlation. For example, increases in elevation may correlate with lower temperatures and therefore be used as a proxy for temperature. Despite the increased robustness and wider applicability of a model incorporating proximal predictors, they are not often included, as distal variables are typically easier to measure (Franklin 2010a). That said, when possible authors should always avoid distal variables as the response a species has to them is more complex. Determining the true variable importance from indirect predictors relies on considerable knowledge of the underlying ecological interactions in a given study area, and must be restricted to a limited geographical extent (as a distal predictor can represent different proximal factors in different regions) (Guisan & Zimmermann 2000). Fortunately, within this study the geographical extent is small, and as such the level of knowledge behind the ecological processes is well known. Given the number of distal predictors incorporated at both sites (Table 4.1), this justification is important if accurate ranking of the variable importance is to be achieved. Furthermore, understanding the proximal/distal nature of the predictors utilised also has an important consideration when attempting to hypothesize the expected shape of the species response curves (Franklin 2010a).

4.2.1.3 Setting multiple working hypotheses

Whittaker (1956) developed key concepts relating to the analysis and interpretation of species distribution abundance along environmental gradients (factors). These response curves are known as ‘species response functions’ for plants, and ‘resource selection functions’ for animals (Austin 2002). Hypothesising the expected shape of these species’ response curves is crucial for several reasons. Firstly, it allows the user’s knowledge and expectations of the species responses to be compared with the output of the models, thereby facilitating clear assessment of each response and whether they

were to be expected (Guisan & Thuiller 2005). Secondly, it allows the responses to be considered in relation to the proposed spatial scale deployed. That is, the extent and resolution of the study area will influence the shape of the response curve. This response could be truncated if the species range extends beyond the limits of the gradient sampled (Austin 2002), or even delayed if extent and resolution are at such a fine scale that there is not sufficient distance from a predictor for a response to be initially registered. Finally, the expected shape of the response curve can influence the type of model utilised. This is because the statistical model attempts to reproduce and formalise the response curve shape (Guisan & Zimmermann 2000). Certain methods (such as generalised linear models (GLMs)) may therefore require transformation of variables, inclusion of polynomial (higher order) terms, and/or interactions between candidate predictors in order to properly fit the predicted response (Franklin 2010a).

The expected shapes of the response curves has been subject to much debate throughout the years (Oksanen 1997; Austin 2002; Oksanen & Minchin 2002). However, the general expectation is that responses to proximal (direct) predictors are either linear or unimodal in shape (Franklin 2010a). Further to this, it is often anticipated that expected unimodal (e.g. bell shaped curves) responses are in fact skewed or irregular (Austin 2002), given that species are influenced by a range of regulatory factors (e.g. biotic and abiotic interactions) (Merow *et al.* 2014). Responses to distal (indirect) predictors are not easily anticipated, and often take complex forms. This form will be depend on the relationship between the indirect predictor and the 'real' underlying causal factor (Franklin 2010a). This does not prevent prediction in response shape to be estimated, however there should be strong ecological correlation between the distal predictor and proximal factors, so causation can be attributed (Austin 2002). In the event of multiple proximal factors being related to a single distal predictor, then the limitations in interpreting causation is likely to be greater and should be acknowledged when interpreting the results.

Figure 4.2 & Figure 4.3 outline the conceptual models for both sites within this study. It should be noted that they do not represent all possible environmental interactions and mechanisms, as this would be complex in design and beyond the scope of the current study (but could be achieved by taking a more mechanistic approach in future). Instead they present a concise explanation of the reasoned relationship between distal predictors and proximal factors, as well as the mechanisms (either physical or environmental) that link them. With regard to the terminology used, whilst 'nutrition' and 'water' are obvious proximal factors that affect animal physiology, the term 'environmental stress' was used as a broad hypernym to encompass climatic (e.g. temperature) and physiological pressures (e.g. from a steep slope) that would be expected to alter sheep behaviour, primarily to conserve energy (and thereby increase the chances of productivity and survivability) (Bird, Lynch & Obst 1984).

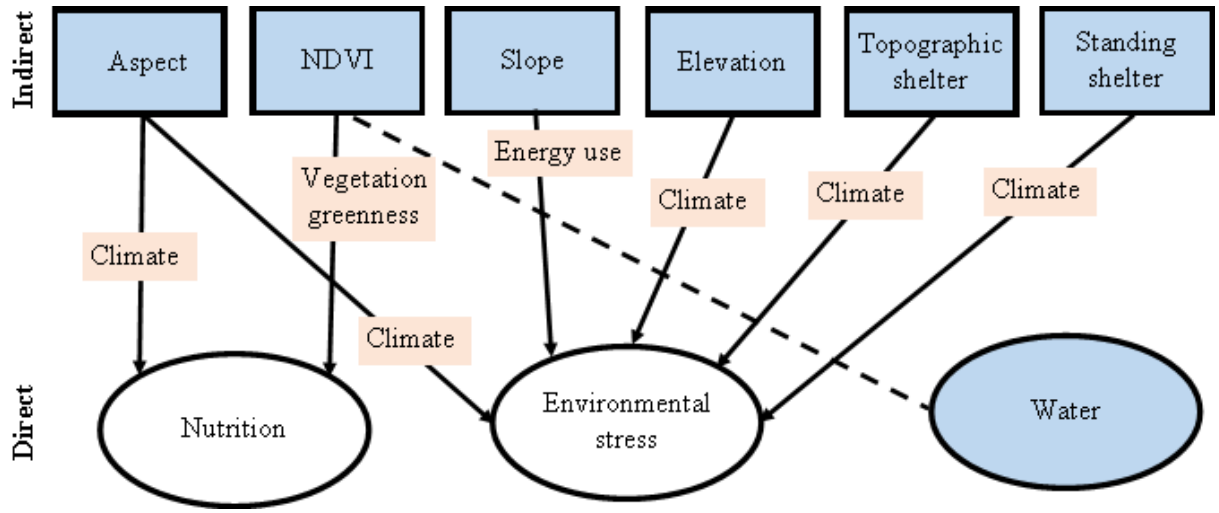


Figure 4.2 Conceptual model for Ffridd Fawr site. Blue fill = measured predictors. Square boxes = indirect (distal) predictors, oval shape = direct (proximal) factors. Orange boxes = linking factors. Dashed line = Possible explanatory link. Adapted from Guisan and Zimmermann (2000).

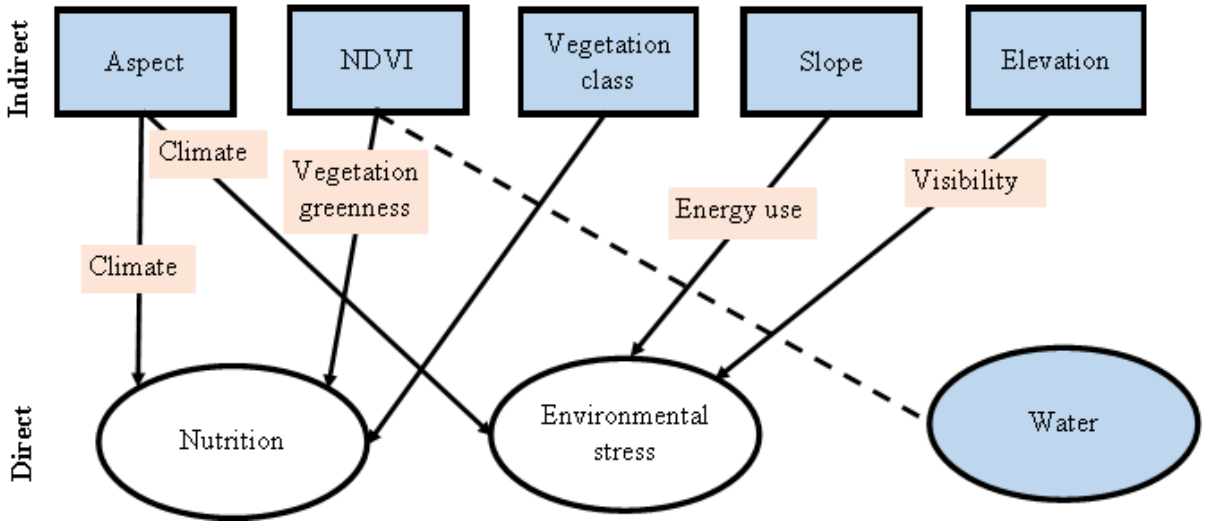


Figure 4.3 Conceptual model for Penglaneinon site. Blue fill = measured predictors. Square boxes = indirect (distal) predictors, oval shape = direct (proximal) factors. Orange boxes = linking factors. Dashed line = Possible explanatory link. Adapted from Guisan and Zimmermann (2000).

The predicted response shape for water was thought to be linear, as it was expected that probability of occurrence would likely decrease as distance to water increased. This was also the case for NDVI, with an expectation that occurrence would increase as NDVI values increased. Though the main utilised correlation for NDVI was nutrition via vegetation greenness (Pettorelli *et al.* 2005; Austin 2007), a secondary possible relationship with water was also acknowledged as others have utilised it in the literature in this way (Feilhauer, He & Rocchini 2012). Aspect (azimuth scale = 0 - 360°) at both sites was predicted to produce a unimodal response, with the expectation that south facing slopes (azimuth = 180°) would be more desirable, as longer sun exposure would result in warmer, more

favourable climatic conditions and also more nutritious vegetation. This is a broad assumption however, and as such it was acknowledged that a skewed response could well be probable. A linear response for 'slope' was expected at both sites, with 'energy use' being considered as the linking mechanism, though for slightly different reasons. Given the steep gradient of the slope at Ffridd Fawr, it was expected that sheep would avoid steep slopes to conserve energy. However, at Penglaneinon, though the gradient of the slopes was more gradual, it was reasoned that an association with thick, tussocked *Molinia* would exacerbate the energy required to navigate the slopes, and also be more challenging with regard to maintaining balance, thus affecting energy conservation in a different manner. Elevation was expected to be linked to environmental stress by climate at Ffridd Fawr, as the lowest part of the site is in a valley bottom, and the top is at the summit of the hill. The expectation therefore was that climatic conditions at the summit would be more environmentally challenging given the increased level of exposure, and that sheep would seek lower elevations. At Penglaneinon, the opposite response was expected. Observations during early site visits noted that visibility was reduced for much of the site, with a clear viewshed of the surrounding area only being visible at higher elevations. The nature of the site topology, combined with the dominant thick homogenous properties of the *Molinia* expanse, make landscape features difficult to discern. Given these observations were made at human eye line, the expectation was that this effect would be exacerbated for the sheep. It was therefore theorised that the sheep would be drawn to elevated parts of the site as a means of improving their outlook. This would both be advantageous with regard to resource identification (e.g. water), but also as a latent predatory response (i.e. to be able to see oncoming predators) (Bangs *et al.* 2005). Given the degree of shelter offered by the *Molinia*, the shallow gradient of the elevation rise, as well as weather protection offered by the surrounding topology, it was not expected that climatic affects (as theorised at Ffridd Fawr) would be notable at Penglaneinon at higher elevations. At both sites, the response curve shape to elevation was expected to be linear (though in opposing directions). At Ffridd Fawr, a linear response was also estimated for standing shelter, with the expectation that probability of occurrence would decrease as distance to standing shelter increased. Given topographic shelter was calculated by azimuth values, with highest shelter being offered at 1° and lowest at 236.9 °, it was also expected that the occurrence would decrease linearly. The clear identified linking mechanism for both shelter types is climate (Bird, Lynch & Obst 1984). Finally, vegetation class at Penglaneinon acted as a proxy for nutrition. Given the categorical nature of the predictor and individualistic manner of the classes (i.e. not being on an environmental gradient), an expected response curve shape could not be predicted. Instead, the only expectation was that the class 'mixed grass species' would have the highest probability of occurrence (as it would likely be most nutritionally diverse).

A critique of these conceptual models is that many of the distal factors are mechanistically linked by climatic variables to the proximal factors. The argument would be that these climatic variables should

be measured instead as they have a more direct link to the physiological responses. It was acknowledged in Chapter 2 that measuring climatic variables would indeed be prudent, but because of resource availability was not considered feasible. It was however suggested as a consideration that should be undertaken in future research. Though this creates limitations in these current models, it should be noted that even in the event of being able to measure the climatic variables it would still be beneficial to include these distal predictors. Primarily, this is because the climatic variables are unable to be manipulated directly, and only through influenceable predictors such as shelter can this be achieved. Therefore, using the current distal predictors alone allows sufficient insight into the potential effect manipulation techniques would have. However, it is acknowledged that complementing these predictors with the incorporation of climatic predictors in future would permit a more comprehensive understanding.

Table 4.1 Summarised conceptual model characteristics and response hypothesis for both sites.

Site	Predictor variable	Factor type (Proximal or distal)	Underlying proximal factor	Hypothesised response curve shape
Ffridd Fawr	Aspect	Distal	Nutrition, Environmental stress	Unimodal
	Slope	Distal	Environmental stress	Linear
	NDVI	Distal	Nutrition	Linear
	Water	Proximal	N/A	Linear
	Standing shelter	Distal	Environmental stress	Linear
	Topographic shelter	Distal	Environmental stress	Linear
	Elevation	Distal	Environmental stress	Linear
Penglaneinon	Aspect	Distal	Nutrition, Environmental stress	Unimodal
	Slope	Distal	Environmental stress	Linear
	NDVI	Distal	Nutrition	Linear
	Water	Proximal	N/A	Linear
	Vegetation class	Distal	Nutrition	N/A
	Elevation	Distal	Environmental stress	Linear

4.2.2 Data preparation for statistical model formulation

The collection of data, and the reasoning behind the sampling approaches employed has already been explored in detail for both the environmental predictors (Chapter 2) and species distribution data (Chapter 3). However, it is important to recognise the characteristics of the data used within the

context of species distribution modelling. In particular, it is crucial that the data used are prepared properly, so that bias (e.g. positional bias)/error (e.g. inaccurate locations) are not unwittingly introduced to the modelling process (Zimmermann *et al.* 2010).

4.2.2.1 Environmental predictors

The first consideration for the chosen predictors is that they are ecologically meaningful. The reasoning stated in Chapter 2 plus the design the of conceptual model offer justification for this to be true in the case of this study.

The second consideration is that species distribution models operate under the assumption that the predictors are independent from each other (i.e. not correlated) (Dormann *et al.* 2013). The presence of correlated predictors (multicollinearity) creates uncertainty in the interpretation of predictor variable importance, as the understanding of animal-environment interactions becomes blurred. Ultimately, this can lead to true causal predictors appearing insignificant (Franklin 2010a). It is therefore important that multicollinearity between predictors is tested before model execution. In the event of a collinearity occurring then the affected predictors can either be left out prior to model selection, or be transformed in relation to each other (Leathwick *et al.* 2005).

Another preparatory step is to ensure that all the environmental variables are of the same geographic projection/coordinate reference system (CRS), and are in perfect alignment to each, both in extent and resolution. Given the raster format of the data, this is necessary so that values for each predictor correspond for every individual pixel.

The next consideration is that the data collection of the predictors (in this case, the UAV imagery) should temporally match the collection time period of the species distribution data (i.e. GNSS points). The reason for this is so that the predictors used in the model represent the true environmental conditions of the animals when their locations were recorded. As noted in Chapter 2, the development issues of the UAV delayed the first successful mission at Ffridd Fawr a month beyond the end of the tracking period, whilst the one survey which best encompassed the tracking period at Penglaneinon was selected. It has already been stated that this poses limitations in the design which should be acknowledged when interpreting the model outputs. Consequently, it is the assessment of the methodological approach, which is the primary consideration in this study, with the specific results only being used to validate the methods employed. After all, true biological inference is only necessary when manipulation is occurring. Given this manipulation would need to be undertaken dynamically within a season since the variable importance for animals one year may not be applicable the following year. Thus, the biological results from this study may carry little relevance beyond validating the methodology employed.

The final consideration for the predictors is the resolution used for the model. In SDM terminology, the scale of the predictors refers both to their grain (i.e. pixel size) and extent (Guisan & Thuiller 2005). Given the extent of both sites in this study is set by the physical boundary (i.e. fence) imposed, only the grain can be altered. Though a 0.5 m ground sampling distance (GSD) was set for the predictors in Chapter 2, this level of data will likely be computationally expensive. Though this is a practical concern that varies with the computational resources available to a specific study, it is acknowledged in the literature as a key element in model design (Renner & Warton 2013). This factor has the ability to restrict the number of different models employed or iterations completed, thereby influencing model output accuracy and validity. Furthermore, it is unlikely that animal response to a predictor will vary much (if at all) over such a fine grain, and therefore it may well not be beneficial in achieving improved biological inference. A more critical aspect to determining the chosen resolution is the positional certainty expressed in the species distribution data. This is important because the distribution data overlays predictors along the exact coordinates specified. Therefore any uncertainty in the position of the points could lead to incorrect pixel values being used, thereby leading to inaccurate estimation of the species–environment relationship (Naimi *et al.* 2014). To understand the level of uncertainty likely, the characteristics of the species data must be explored and understood.

4.2.2.2 The species data

Species distribution data in SDMs can be subdivided into two forms: presence data and absence data. Presence data corresponds to a positive sighting of an individual from the study species at a particular point in space and time. Absence data refers to observations of where an individual has not been witnessed (Pulliam 2000). Crudely stated, presence data are thought to inform about the locations that are environmentally suitable for a species, whereas absences inform on the environments that are not suitable. In reality, locations where absences are observed can in fact be environmentally favourable, and instead factors such as dispersal limitations, historical factors, local extinctions, biotic interactions, and patch size may be responsible (Lobo, Jiménez-Valverde & Hortal 2010). Ideally, both presence and absence data should be incorporated into a species distribution model, as it typically represents a more systematic survey of a site. However, in many cases it may only be possible to collect presence data (e.g. in tracking studies) (Renner & Warton 2013). Given most methods assume the input of both presence and absence data, it is necessary to produce either pseudo-absence or background data when only presence data is available (Franklin 2010a). Because GNSS tracking was used to record animal locations in this study, only presence was collected. Therefore, it will be necessary for either pseudo-absence/background data to be generated depending on the particular models being used.

Aside from presence/absence nature of the data, sample size is also an important consideration. Sample size is positively related to the performance of species distribution models (Hirzel & Guisan 2002), with reported sample sizes of ~30-100 observations thought to be sufficient to achieve acceptable model performance (Stockwell & Peterson 2002; Elith* *et al.* 2006). In their conclusions, Coudun and Gégout (2006) proposed a general rule of a minimum 50 occurrences for modelling when they quantified the sensitivity of response curves to sampling characteristics. The number of observations recorded for both sites in this study (following data cleaning) were 366 (Ffridd Fawr) and 522 (Penglaneinon), and as such are sufficient in this regard.

When modelling species distributions with presence-only data in a presence/background or presence pseudo absence design, a critical assumption is that location data has been collected without bias (Stolar & Nielsen 2015). This bias may be temporal (e.g. collecting data at sporadic time intervals without consistency) or spatial (e.g. not covering a site systematically, recording observations next to a road for instance) in nature (Bean, Stafford & Brashares 2012). Given GNSS tracking was utilised in this study to locate the animals, sampling was not temporally biased as the recordings were set at even 8 days intervals and recorded for a whole day (24 hr) at a time. Therefore, although there was a weekly gap between recordings, it was consistent throughout. Sampling was also not spatially biased, as locations were recorded at the true location (within the realms of GNSS error) of an animal and not subject to observer position, visibility or availability (Boria *et al.* 2014). As such detectability of locations points was maximised, with omissions expected to be minimal or non-existent (especially given the sampling strategy employed; see below). Furthermore, given a number of animals were collared (>5 at each site), there is certainty that locations observed belong to a representative sample of the individuals, whereas in conventional observer recordings, the number of individuals will often be unknown. As such, data preparation techniques such as spatial rarefication or bias file creation which are used to alleviate bias effects in the model, were not required (Brown 2014; McCoshum *et al.* 2016).

The final consideration, which was alluded to previously (with regard to the effect on predictor resolution) is the likely positional uncertainty of the species data. As explained before, understanding the likely degree of error is crucial in order not to introduce uncertainty in the model outputs (Moudrý & Šímová 2012). The obvious point when positional uncertainty will cause likely inaccuracies, is when the level of uncertainty (accuracy) in the species data exceeds that of the predictor resolution. This is because it cannot be guaranteed in such a circumstance that the values belonging to an individual location point definitively correspond with a specific pixel in the predictor stack. Generally, the high level of accuracy offered by modern GNSS tracking systems (~10m) means that the positional uncertainty is sufficiently below the resolution of the predictors employed. However, the 0.5m GSD resolution available for the predictors in this study is unprecedented (or certainly rare)

within the SDM literature, and far below the uncertainty level of the species data. Although the specific accuracy of collars employed in this study could regrettably not be precisely measured (COVID-19 restrictions prevented the necessary validation field work), knowing the particular GNSS module utilised (Quectel L86) means that the stated manufacturer accuracy is known (<2.5 m circular error probable (CEP), 95% radius (R95) = 4.75 m (using crude assumption of (2.5/50(CEP)))*95(R95)). Though the testing conditions for this claim are unknown, by extracting the animal location points from the middle of high intensity recording bursts in the data of this study, the effective accuracy will maximised. This is based on the assumption that GNSS accuracy will typically be greater at high (e.g. 1 s) rather than low intervals (e.g. 1 h) (Yousef & Ragheb 2014; Forin-Wiart *et al.* 2015). Unlike Johnson & Gillingham (2008), it was also deemed that introduced error in the GNSS accuracy, such as through decreased line of sight (e.g. because of trees, topography) was unlikely to be a concern. Therefore, it would be reasonable to assume that the accuracy of the recordings used was likely to be ~5 m (R95), but certainly under <10 m (R95). Regardless of this, it is still evident that the positional uncertainty will be surpass the starting resolution of the predictors at both sites, and as such remedial action will be required to account for this uncertainty. One such approach developed by Naimi *et al* (2014) used the k-statistic to quantify the local spatial association in the predictors at each species occurrence location, followed by a probabilistic Monte Carlo simulation to assess the effect of positional uncertainty in location data on an SDMs performance. Another more simple approach, outlined by Johnson & Gillingham (2008) is to aggregate pixels until the error threshold of a species location is confined within a single patch (cell). Given that at such a fine grain, predictor response would unlikely vary much (if at all), as such it would be reasonable to suggest that aggregation may prove sufficiently effective. A proposed solution would therefore be to execute the model using predictors at their native resolution (0.5 m) and compare the variable importance outputs with predictors aggregated beyond the likely positional uncertainty in the species locations (10 m GSD to be confident). This would allow an examination of the collective effect positional uncertainty was having on the outputs and inform as to whether this uncertainty compromised biological inference of the results if consensus in predictor ranking (for variable importance) did not exist.

A summary of all considerations discussed for the data preparation stage of the species distribution modelling procedure is viewable in Table 4.2

Table 4.2 A summary of all considerations discussed for the data preparation stage of the species distribution modelling procedure. Potential problems, and the justification for mitigating or not are surmised.

Consideration	**Possible problem**	**Action required?**	**Justification/mitigation**
Ecologically meaningful?	Model outputs incorrectly infer true ecological processes.	No	Detailed reasoning and justification in Chapter 2 and conceptual model.
Multicollinearity	Understanding of animal-environment interaction becomes blurred.	Yes	Perform multi-collinearity test. Remove/transform affected predictors.
Matching CRS and alignment in predictors	Models cannot be executed without alignment.	Yes	Check CRS and alignment of predictor before model execution.
Temporal matching with species data	Predictors not representing true environmental conditions when species data was collected.	No	Though a perfect temporal overlap does not exist between data, it is sufficient to address the study objective (methodological validation.
Predictor resolution	Native resolution too computationally expensive to model, potential issues with positional uncertainty	Yes	See positional uncertainty.
Species data type	No absence data available.	Yes	Produce pseudo-absence or background data depending on model used.
Insufficient sample size	Insufficient sample size for either site negatively affecting model performance.	No	Sufficient number of observations at both sites (522 & 366).
Sample bias	Spatial or temporal bias influencing model output.	No	Scheduled GNSS recording ensured even sampling. No observer caused bias because of the method employed.
Positional uncertainty	Uncertainty in model outputs.	yes	Compare variable importance from native predictor resolution

	(0.5m) and aggregated predictors (10m) to assess uncertainty effect.

4.2.3 Model selection and fitting

Now that the conceptual model has been defined, and the characteristics and limitations of the data available explored, it is now prudent to select the model(s) that best suit the expected response variables and data present (Guisan & Zimmermann 2000). Modelling methods are typically either statistical (e.g. GLMs), or based on machine-learning techniques (e.g., maximum entropy (Maxent) and artificial neural networks (ANNs)) (Pearson 2007). Statistical regression-based models are often considered as mostly parametric or model driven approaches to statistical learning, in which assumptions are made on the form of the model, and which use all data to estimate the parameters (global method). In contrast, machine learning methods comprise various algorithms that are used to learn the mapping function/classification rules inductively, directly from the training data (Franklin 2010a).

Many of the modelling methods are remarkably flexible with regard to measurement scale, predictor distribution, and response variables. However, some models contain advantages over others depending on the data types (presence-only, presence/background, presence/pseudo-absence) and predictor response shape. Whilst numerous articles within the literature discuss the various nuances to each model and their suitability to different circumstances (Pearson 2007; Marmion *et al.* 2009; Franklin 2010a; Hao *et al.* 2020), understanding the 'best' model to select is often complicated and not straightforward, with Segurado & Araujo (2004) concluding that it is '*unlikely that a single best habitat modelling procedure will ever be identified'*. After all, model outputs have been shown to be sensitive to steps of the modelling process such as parameterisation or selection (Buisson *et al.* 2010). As such, much attention has been given to the investigation of model based uncertainty, with two main approaches being used to overcome this: (1) judging best performance through multiple model comparisons; and (2) combining models to enable ensemble forecasting (Marmion *et al.* 2009).

Despite ensembles first being introduced into statistical mechanics in 1878 by J. Willard Gibbs, their use in species distribution modelling is a comparatively recent development. In their seminal study of using ensembles within SDM, Araujo & New (2007) defined them in a pure sense as '*an idealization consisting of a large (possibly infinite) number of copies of a system, considered all at once, each of which represents a possible state that the real system might be in at some specified time*'. Put simply, the rationale behind their use is that by combining different model types over the same input data, a range of predictions can be generated thereby allowing individual model uncertainty to be explored (Franklin 2010a). This approach of applying several methods within a consensus framework has been

demonstrated by several studies to substantially improve SDM outputs (Crossman & Bass 2008; Marmion *et al.* 2009; Grenouillet *et al.* 2011). However, it should also be noted that a recent study (Hao *et al.* 2020) has shown that having experienced knowledge (both theoretically and practically) of the particular modelling approach employed and the ability to 'tune' the parameters properly, will likely yield the best predictive performance regardless of whether an ensemble is used or not. That said, they found that ensembles performed well compared to untuned models. Given the main output of interest in this study is the variable importance of the predictors and not the predictive mapping of the SDMs, a prudent approach would be to select a number of different models, compare the variable importance from each model and their level of consensus, then perform an ensemble estimation of the variable importance. Regarding the models that should be selected, the following were chosen based on the characteristics of the data to be analysed and the expected shape of the response curves, with a view to contain representation of both statistical (regression based) and machine learning approaches if possible.

Table 4.3 Modelling method and type to be used in this study.

Model type	**Method**
Statistical (Regression based)	Generalised additive model (GAM)
	Multivariate adaptive regression splines (MARS)
Machine learning	Maximum entropy (MaxEnt)
	Artificial neural networks (ANN)
	Random Forest (RF)
	Boosted regression trees (BRT)
	Support vector machine (SVM)

With regard to the statistical approaches to be utilised, given the response curves were expected to be non-linear for many predictors (i.e. either unimodal or skewed), it was decided that generalised additive models (GAMs) and multivariate adaptive regression splines (MARS) would be best suited. GAMs are a non-parametric extension of generalised linear models (GLMs), and are often preferred over GLMs because of their ability to better deal with highly non-linear relationships between the species response and predictors (Guisan, Edwards Jr & Hastie 2002). MARS is typically viewed as a generalisation of step-wise linear regression (Friedman, Hastie & Tibshirani 2001), where the technique combines the strengths of spline fitting and regression trees by replacing the step functions normally associated with regression trees with piecewise linear basis functions (Leathwick, Elith & Hastie 2006). As a result, like GAMs, MARS is able to model complex relationships between species

response and predictor variables, and has the advantage of being computationally fast to execute (Franklin 2010a).

The use of machine learning modelling methods does not require a pre-defined ecological hypothesis in order to operate successfully, as the algorithms primarily aim to uncover patterns in the data (Gobeyn *et al.* 2019). These data-driven approaches therefore often include fewer assumptions on response curve shape (Gahegan 2003), and are better able to incorporate both continuous and categorical data. As there are a number of supervised machine learning approaches applicable to SDM, it was deemed suitable to select models representing different options, so the uncertainty across the array of approaches could be explored and accounted for. These included decision-tree based methods (random forest (RF) and boosted regression trees (BRT)), artificial neural networks (ANN), maximum entropy (MAXENT), and support vector machines (SVM). A comprehensive and detailed description of each approach and their differences can be found in Franklin (2010a).

4.2.4 Model evaluation

Once the selected models have been fitted (calibrated), it is then necessary to evaluate their respective performance with regard to their predictive accuracy and validity of outputs. Rykiel (1996) noted that a model performance can only be deemed acceptable for its intended use when it meets the specified performance requirements. Predictive accuracy assessment through empirical measurements, ecological realism of the outputs, and model credibility (acceptability to the user community) are all important criteria for evaluating model performance (Franklin 2010a).

The predictive power of a model may be evaluated two ways. The first and most robust method is to use an independent data set to that of the training data to validate. However, in most studies this data is not available, and so a resampling method is adopted (Guisan & Zimmermann 2000). This approach can be achieved through *k*-fold partitioning, jack-knife, bootstrapping techniques, or subsampling. In the absence of an independent species data set for either site in this study, a resampling approach was necessary. Measurement of the accuracy can then be assessed through a number of metrics, of which the most frequently used options include the ‘area under the curve of the receiver-operating characteristic’ (AUC), the true skill statistic (TSS), and the Kappa statistic.

4.2.5 Objectives

The aim of this chapter was to rank the variable importance of environmental predictors at the separate study sites. In order to achieve this, a consensus framework consisting of a series of species distribution modelling algorithms was employed using a combination of statistical and machine-learning methods (Table 4.3). Comparisons between model performance was made, with the most empirically accurate methods being used to form a final ensemble for each site. Predictor resolution of both 0.5 m and 10 m GSD were assessed separately and compared to determine the effect of

positional uncertainty on the outputs. The ecological realism of the outputs, combined with the knowledge of each site, was used to assess model credibility. Finally, the effectiveness of the models, their limitations, and the recommended areas of future work were all discussed.

4.3 Model development

4.3.1 Raster data preparation

All data preparation and modelling was undertaken in R (vers 3.3.6) using Rstudio (vers 1.2.5033). As the 0.5 m rasters had already been resampled and projected to WGS84-30N (for the first chapter), they required no further pre-processing. For 10 m resolution predictors, the 0.5 m predictors were resampled using the bilinear method in the 'resample' function in the R 'raster' library (Hijmans *et al.* 2015). In order to guarantee alignment between predictors, rasters at each resolution were stacked. With regards to the naming of Fridd Fawr predictors, topographic shelter was defined as 'Topo_shelter' and standing shelter simply as 'Shelter'.

4.3.2 Assessment of collinearity

Collinearity (i.e. correlation between two or more predictor variables) was assessed using the Variance Inflation Factor (VIF). VIF is based on the square of the multiple correlation coefficient resulting from regressing a predictor variable against all other predictor variables (Guisan, Thuiller & Zimmermann 2017). Two strategies from the 'usdm' library were used to assess collinearity; 'vifcor' and 'vifstep' (Naimi *et al.* 2014). 'Vifcor' and 'vifstep' both exclude highly collinear variables through a stepwise procedure. 'Vifcor' finds a pair of variables which has the maximum linear correlation ($R > 0.7$) and excludes the one that has the greatest VIF. The procedure is repeated until no variables with a high correlation coefficient ($R > 0.7$) remains. 'Vifstep' calculates VIF for all variables, excludes the one with highest VIF ($R > 0.7$), then repeats the procedure until no variables with a VIF ($R > 0.7$) remains. VIF values of >10 were taken as indicative of problematic collinearity/redundancy (Montgomery, Peck & Vining 2012; Pradhan 2016).

At both sites/resolutions, collinearity was low, and as such no predictors were removed. For Ffridd Fawr, the maximum correlation between predictors was observed between 'topographic shelter' and 'aspect' (*vifcor*, $R = 0.57$) at the 10 m resolution, whilst the largest VIF (also at a 10 m resolution) was for topographic shelter (*vifstep*, VIF = 1.97). For Penglaneinon, the maximum correlation was again at the 10 m resolution (*vifcor*, slope ~ elevation, $R = -0.48$), as was the largest VIF score (*vifcor*, VIF = 1.59 (slope)).

4.3.3 Animal location data preparation

Two-hour interval location data was used for the analyses for the two subsequent reasons. Firstly, in their study using cattle, Pertotto-Baldivieso *et al* (2012) found that auto-correlation between successive observations could be minimised if time intervals were separated by at least 2 h. Secondly, given the obvious range restrictions of sites in this study (by being completely enclosed), it was reasoned that the animals at both sites could reasonably cover the whole range available to them within an hour, and therefore a 2 h interval length allowed for full site coverage between recordings. To minimise positional uncertainty per location, these two-hour intervals were extracted from the 1 s data from the logger schedule's eighth 'intensive' recording day. This was based on the assumption that GNSS accuracy will typically be greater at high (e.g. 1 s) rather than low intervals (e.g. 1 h) (Yousef & Ragheb 2014; Forin-Wiart *et al.* 2015). A total of 12 recordings per 8 d was therefore available for each animal. This resulted in a total of 532 analysed observations for 5 sheep at Penglaneinon across a 105 d period (29/05/18-10/09/18), and 363 observations for 7 sheep at Ffridd Fawr across a 49 d period (29/05/18-16/07/18). Location data was then loaded into R, re-projected from WGS84 into WGS84 30N (to match the rasters), then transformed into a spatial points data frame. Points were then plotted against the rasterstacks to validate alignment.

4.3.4 Model fitting and evaluation

Species distribution modelling was undertaken using the 'sdm' package (vers 1.0.89) (Naimi & Araújo 2016). First, an 'sdmdata' object was created that held species data and explanatory variables. During this stage, one thousand background points (pseudo-absence) were randomly generated for each site, with any points that matched the locations of presence data removed. A seed was set between resolutions so background data was the same. Four models (one for each resolution at both sites) were then fitted and evaluated. A Multilayer perceptron (MLP) and Radial-basis function (RBF) were selected as the artificial neural networks (ANNs)(in accordance with the 'sdm' library options), as well as GAM, MARS, SVM, RF, BRT, and MaxEnt as the other modelling methods. A dual partitioning method of subsampling and bootstrapping was used for evaluation. *K*-fold cross-validation was also explored initially (using 10 folds), however as accuracy results did not alter between methods, it was excluded to save processing time. The partition size for subsampling was set to 29% (Franklin 2010a). The number of iterations for each modelling method was set to 100, as this acted as a compromise between the length of processing time required, and the need to allow sensitivity to changes in parameters to be visualised (Engler, Hordijk & Guisan 2012). A seed was also set for model fitting between resolutions, so training data was the same. Parallel processing using in-built functionality in the 'sdm' library as well the 'cluster' function in the 'raster' library was incorporated where possible, particularly in the sdmdata object creation, model fitting, and evaluation stages to speed up processing time (35 cores were utilised).

4.3.5 Calculating variable importance and plotting response curves

Variable importance for all four models was calculated using the 'getVarImp' function in 'sdm', then converted into a dataframe to facilitate further analysis. Within the 'getVarImp' function, Pearson's correlation was chosen as the metric by which variable importance would be expressed (Thuiller *et al.* 2009). Training and test data were both used to delineate variable importance. Given the inability to measure the predictors when animals were being tracked, caution had to be observed when inferring the ecological implications of the SDM outputs. As such, detailed statistical comparison (e.g. analysis of variance) between predictors in importance rankings were avoided so as not to overstate results. Instead, the frequency of each predictor at every importance level was calculated to better assess agreement between iterations. Variability between model methods for permutation importance of each predictor was also assessed, as was their individual rankings of importance. Response curves were plotted, and their interpretation compared with the hypothesised responses. Due to confirmed library issues involving the 'getVarImp' function, categorical data, and ANNs, the response curves could not be generated for a proportion of iterations at Penglaneinon. As such, models were regenerated for the response curves without the vegetation classification. There was no visible difference between the response curves of the models for the remaining predictors.

4.4 Model analysis and discussion- Fridd Fawr

4.4.1 SDM predictive accuracy at different resolutions

Using AUC values as the main proxy for accuracy, the predictive accuracy of models at the two different predictor resolutions was observed to be similar, with the increases in performance for 10 m models (M = 0.795) being slight (+0.015 AUC) against 0.5 m models (M= 0.78) (Table 4.4). Predictive accuracy for all models was moderate in performance (0.7-0.9 AUC). At both resolutions, Random Forest (RF) was estimated to be the most accurate, and Radial-Basis Function (RBF) the least. Finally, there did not appear to be any notable difference in predictive accuracy between the different modelling method types (i.e. classical regression (M = 0.783 AUC) vs machine learning (M = 0.789 AUC)). Given the level of similarity between model accuracy, all models were included in the ensemble to calculate variable importance.

Table 4.4 Accuracy assessment of different species distribution models at the Ffridd Fawr site. AUC = Area under the curve, COR = Pearson's correlation, TSS = True skill statistics, Deviance = Deviance of the predicted values from the observed values. (±values) = difference in AUC for 10m scores relative to 0.5 m AUC scores. For clarity, both MLP and RBF are different ANN methods.

Resolution	Model	AUC	COR	TSS	Deviance
0.5 m	GAM	0.77	0.41	0.43	0.97
	MARS	0.77	0.41	0.44	0.98
	SVM	0.77	0.39	0.44	1.32
	RF	0.85	0.57	0.56	0.80
	BRT	0.79	0.44	0.45	1.01
	MaxEnt	0.78	0.42	0.45	1.17
	MLP	0.77	0.41	0.45	1.58
	RBF	0.74	0.37	0.41	1.02
10 m	GAM	0.80 (+0.03)	0.46	0.47	0.93
	MARS	0.79 (+0.02)	0.44	0.46	0.97
	SVM	0.78 (+0.01)	0.42	0.45	1.27
	RF	0.85	0.58	0.56	0.80
	BRT	0.79	0.46	0.46	1
	MaxEnt	0.79 (+0.01)	0.45	0.47	1.1
	MLP	0.79 (+0.02)	0.44	0.48	1.4
	RBF	0.77 (+0.03)	0.41	0.43	0.98

4.4.2 Variable importance at different resolutions.

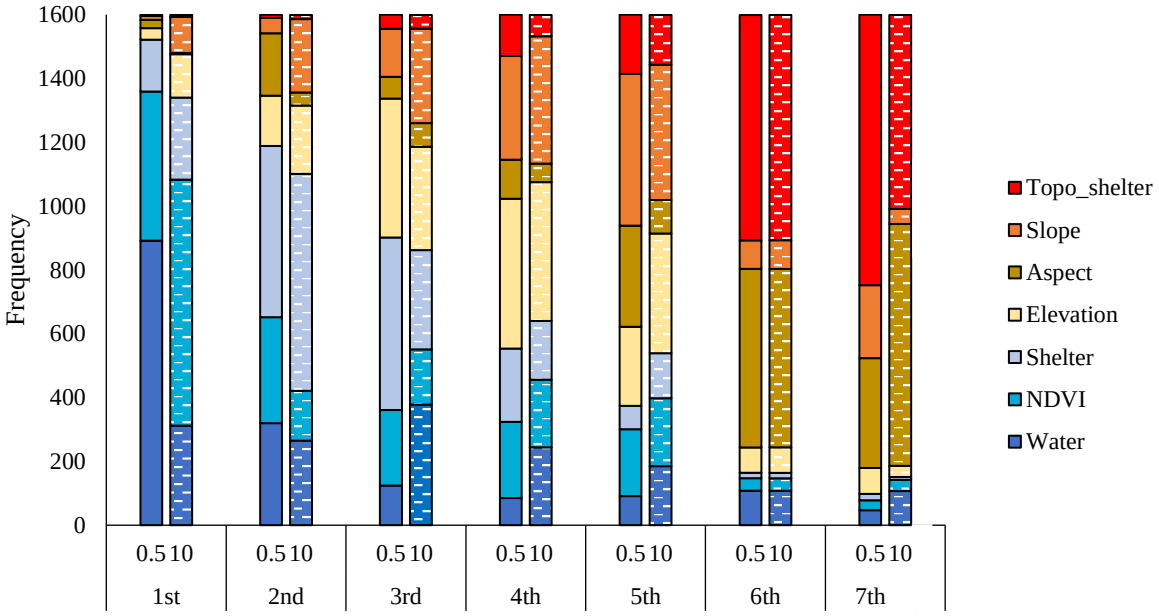


Figure 4.4 Frequency of occurrence for predictors at each level of the variable importance order according to resolution at Ffridd Fawr. Total number of model iterations for each resolution = 1600. Solid bars denote variables at 0.5 m resolution. Hatched bars denote variables at 10 m resolution.

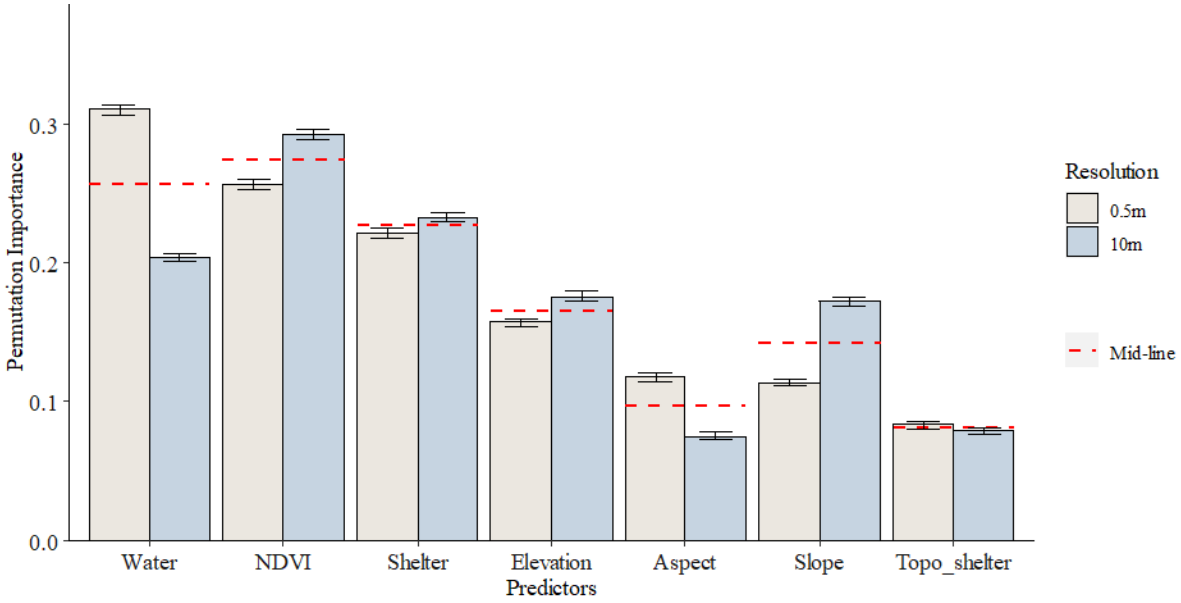


Figure 4.5 Mean variable importance across the 8 different models using different predictor resolutions at Ffridd Fawr. Predictors are ordered on the x axis according to the 0.5 m importance rank to facilitate comparison. Error bars = standard error of the mean (SEM).

It is notable from Figure 4.4 that the order of variable importance alters between the different input predictor resolutions, suggesting that positional uncertainty from GNSS error would likely affect the interpretation of the results. Though the level of uncertainty varies according to the predictor, with 'water' and 'slope' being considerably more affected then 'NDVI' and 'shelter', it is evident that

attempting to discern the true exact order of importance for the predictors is not possible given this uncertainty. That said, Figure 4.4 does indicate potential groupings of importance for various predictors. For example, neither 'aspect' or 'topographic shelter' are considered important variables at either resolution, whilst both 'standing shelter' and 'NDVI' are consistently high in the level of importance. This is further explored and emphasised in Figure 4.5**.** Across both resolutions, 'NDVI', 'standing shelter', and 'water' predominate the frequency of occurrence for the top three most important predictors across all model iterations (n = 1600). In combination, they account for 95% (0.5 m) and 84% (10 m) of the 1st ranking, 74% (0.5 m) and 69% (1 0m) for the 2nd position, and 56% (0.5 m) and 54% (10 m) for 3rd position. 'Topographic shelter' and 'aspect' are consistently the least two important variables (6th and 7th) throughout the iterations, whilst 'elevation' and 'slope' occupy the middle rankings (4th and 5th) most consistently.

When assessing the consensus between the modelling approaches with regard to their predictions of variable importance, it is clear from Figure 4.6 and Figure 4.7 that the level of permutation importance assigned to each predictor varies considerably between model types for both resolutions. The Multi-Layer Perceptron (MLP) models in particular have attached more importance to each predictor across both resolutions then other models. It is also notable that the large variations in variable importance for 'water' and 'slope' between the differing resolutions is attributable to value changes across all the models as opposed to large fluctuations over single model types responses.

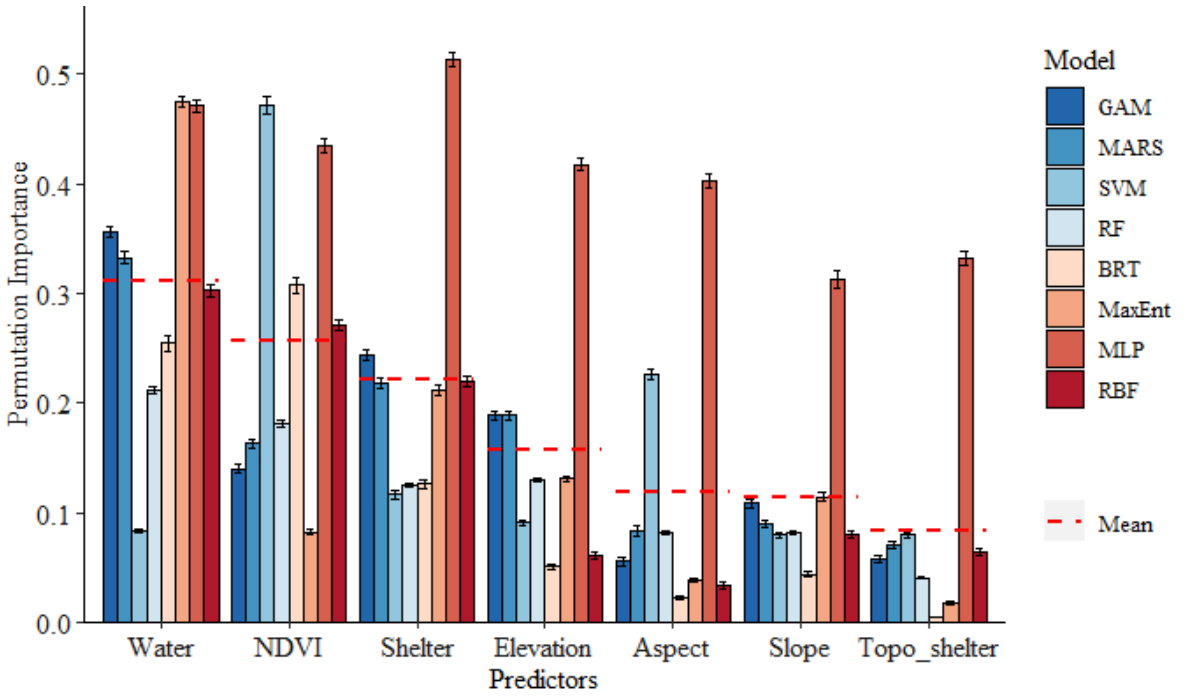


Figure 4.6 Mean variable importance for each predictor according to individual model types at 0.5 m predictor resolution at Ffridd Fawr. Each model is averaged across 200 iterations. Mean line = average across all models. Predictor order is arranged highest to lowest according to averaged variable importance.

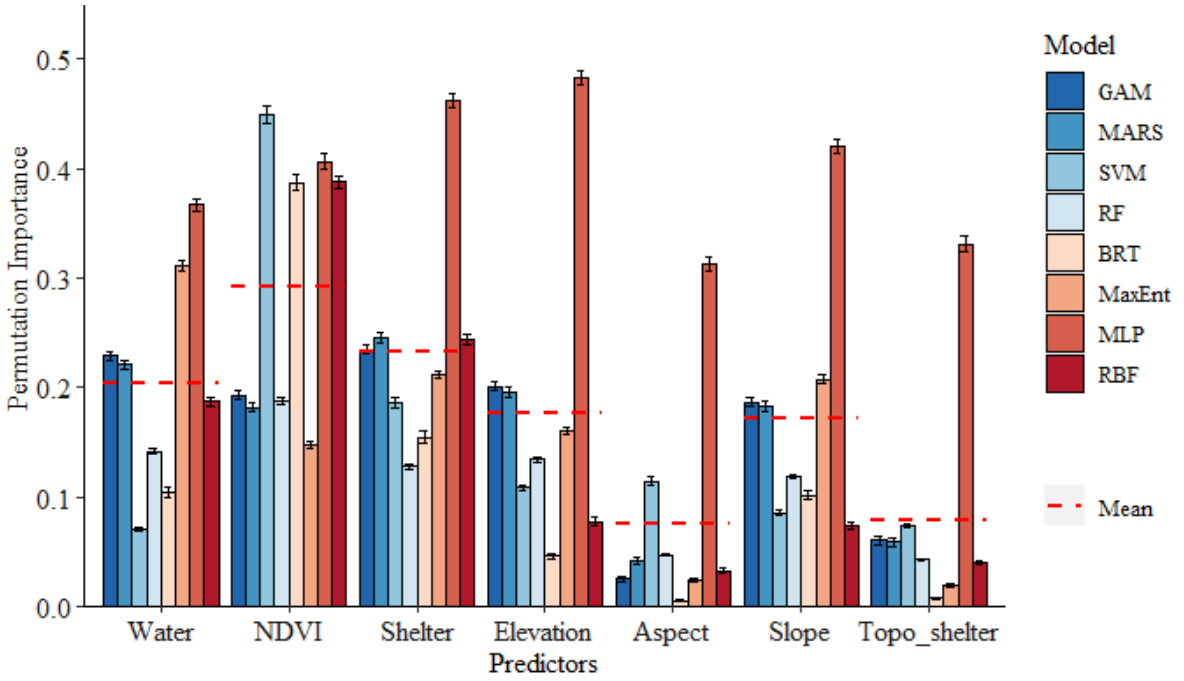


Figure 4.7 Mean variable importance for each predictor according to individual model types at 10 m predictor resolution at Ffridd Fawr. Each model is averaged across 200 iterations. Mean line = average across all models. Predictor order is arranged highest to lowest according to averaged variable importance of the 0.5 m resolution models to facilitate clear comparison.

Table 4.5 Predictor importance ranking at Ffridd Fawr according to each individual model type at both predictor resolutions. * = where two predictors contain the same permutation importance and therefore same rank. For clarity, both MLP and RBF are different ANN methods.

Resolution	Models	Predictors (ordered according to average variable importance)						
		Water (1st)	**NDVI (2nd)**	**Shelter (3rd)**	**Elevation (4th)**	**Aspect (5th)**	**Slope (6th)**	**Topo Shelter (7th)**
0.5m	GAM	1	4	2	3	7	5	6
	MARS	1	4	2	3	6	5	7
	SVM	5	1	3	4	2	6*	6*
	RF	1	2	4	3	5*	5*	7
	BRT	2	1	3	4	6	5	7
	MaxEnt	1	5	2	3	6	4	7
	MLP	2	3	1	4	5	7	6
	RBF	1	2	3	6	7	4	5
Resolution	**Models**	**Predictors (ordered according to average variable importance)**						
		NDVI (1st)	**Shelter (2nd)**	**Water (3rd)**	**Elevation (4th)**	**Slope (5th)**	**Topo Shelter (6th)**	**Aspect (7th)**
10m	GAM	4	1	2	3	5	6	7
	MARS	5	1	2	3	4	6	7
	SVM	1	2	7	4	5	6	3
	RF	1	4	2	3	5	7	6
	BRT	1	2	3	5	4	6	7
	MaxEnt	5	2	1	4	3	7	6
	MLP	4	2	5	1	3	6	7
	RBF	1	2	3	4	5	6	7

When viewing Table 4.5, the variability between model types with regard to variable importance is further emphasised. For the 0.5 m models, only RF agreed with the mean variable importance rank across all the models. The same was also true for RBF in the 10 m models. Further to this, no models at either resolution were in precise agreement with each other with regards to the order.

4.4.3 Response curves

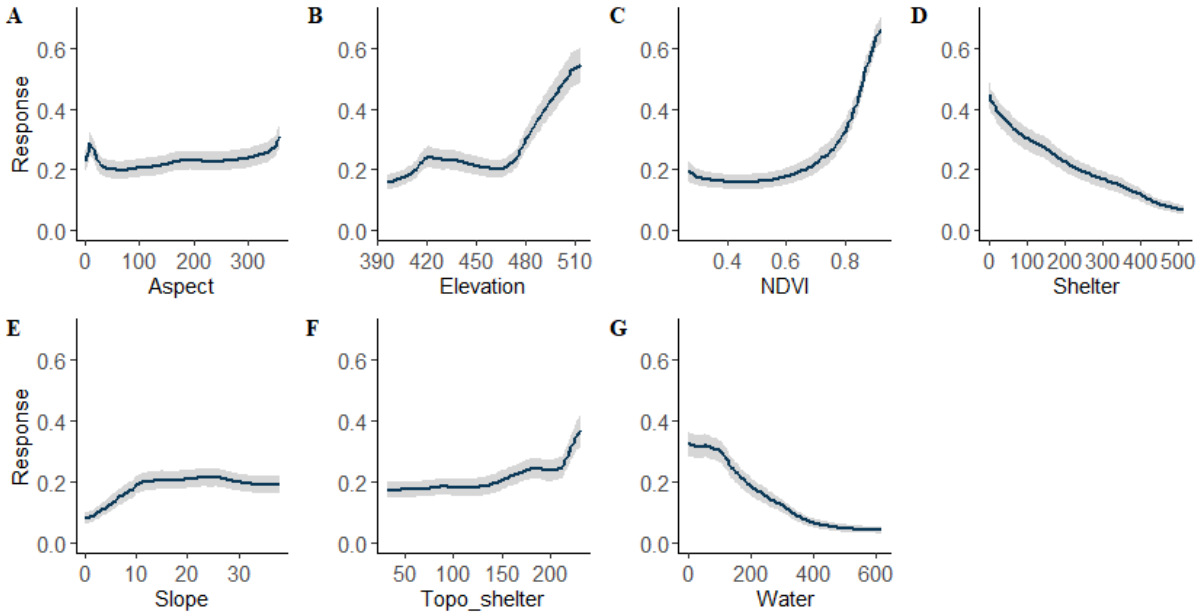


Figure 4.8 Mean response curves for all model types for each predictor at Ffridd Fawr using 0.5 m resolution predictors. Shaded areas denote standard error of the mean.

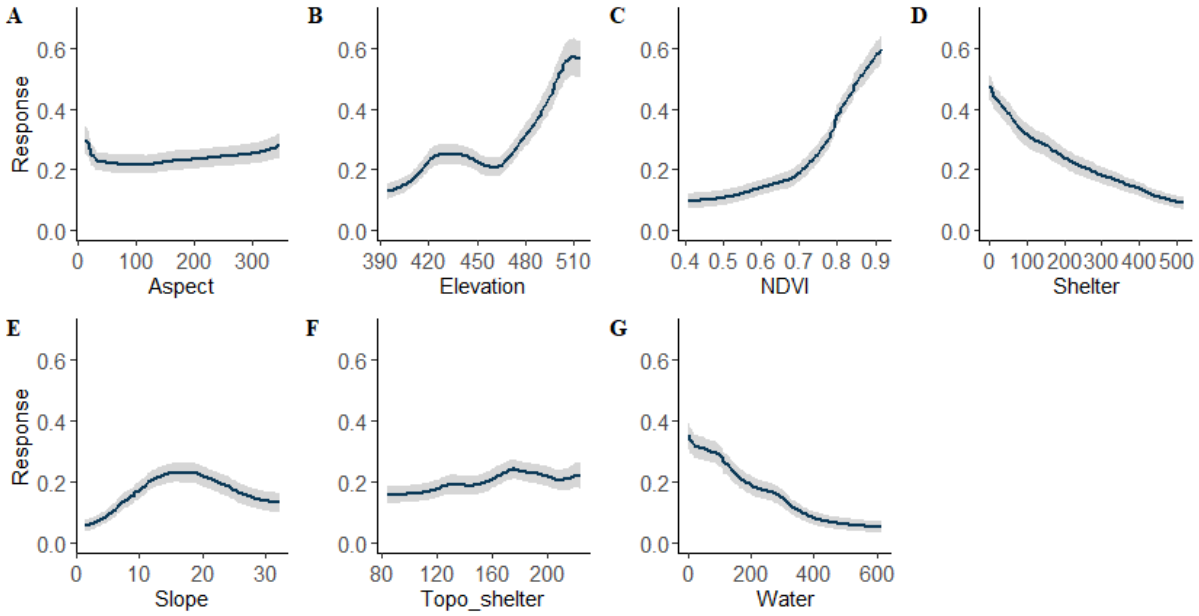


Figure 4.9 Mean response curves for all model types for each predictor at Ffridd Fawr using 10 m resolution predictors. Shaded areas denote standard error of the mean.

As demonstrated by Figure 4.8 and Figure 4.9, the general relationship in the response curves (rcurve) between both resolutions for all predictors was consistent. Though minor variations may be witnessed in the exact form of the response curves, the predominate pattern of responses are markedly similar. This degree of agreement facilitates the assessment of the individual hypothesis relating to the expected response curve shape, as well as the exploration and interpretation of the ecological reasoning behind these forms.

Species response across both resolutions did not appear to change with differences in 'aspect', indicating it had little or no importance to influencing distribution. The expected shape in the event of a response was hypothesised to be unimodal (individuals were thought to show more preference to south facing slopes (180°)), however this was recognised as a tentative and broad assumption, and therefore the lack of response was not unexpected. The observed response for 'elevation' however was not predicted. It was hypothesised that the response would be linear, with lower elevations (containing more favourable climatic conditions) seeing a greater response then higher elevations (where climatic conditions would likely be worse). However, though the observed response was linear, it paradoxically indicated an increased response as elevation increased. The form of the response was also skewed at both resolutions, with the most notable response occurring between altitudes of ~470-500 m. When this incline was thresholded in QGIS, it was evident that lack of points towards the southern end of site was key to the observed response (Figure 4.10). The prominent presence of the animal locations at higher elevations suggests that climate is not affecting distribution as expected, however this requires further research to explore and quantify.

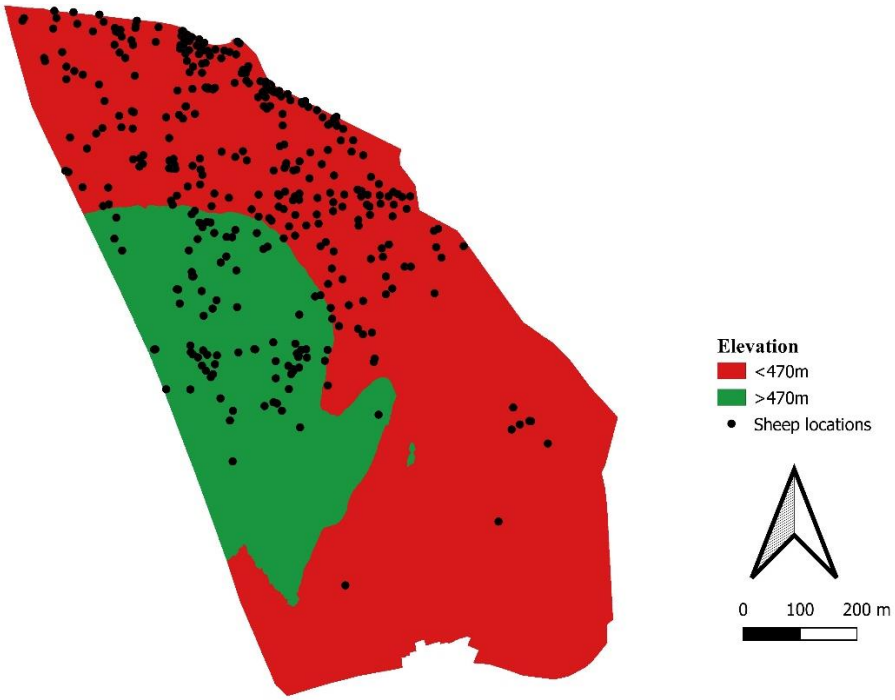


Figure 4.10 Elevation at 0.5 m for Ffridd Fawr segmented into two classes along a 470 m threshold. Green areas denote areas of increasing high response, whereas red areas denote a static response.

NDVI follows the hypothesised linear response, where the response increased as NDVI values increased. The response was more delayed at the 0.5 m resolution with a little change being observed until NDVI = 0.6, whereas in the 10 m resolution curve the response slope gradually inclines to ~0.6 before a steep rise. It is likely that the reduction in spectral heterogeneity between 0.5 m and 10 m resolutions caused the response to appear more smoothed at 10 m resolution. Of all the predictors, the response to 'shelter' was most in agreement with the hypothesised shape. The response decreases in a

true linear fashion the further from a shelter source an animal goes and is true for both resolutions. In contrast, the observed response for ‘slope’ is unimodal, where a linear decrease was expected. That said, the level of response is small (~0.15 - 0.25°), and not consistent between resolutions. As such, it was deemed prudent to not interpret ecological causes to the response when it so small. The response for ‘topographic shelter’ was also small, with an approximate flat form observed. There does appear to be an early response in the 0.5 m plot >220°, however this is not echoed in the 10 m response. Finally, despite the differences in variable importance measurement, the consensus for ‘water’ between resolutions is remarkably similar. The response curve matches the hypothesised outcome, with a linear decrease in response occurring as distance to water increased. There are minor fluctuations in the form of the 10 m response which may explain part of the variation in variable importance between the resolutions, however the general agreement between them suggests water to have a high degree of importance regardless of resolution.

4.5 Model analysis and discussion- Penglaneinon

4.5.1 SDM predictive accuracy at different resolutions

Table 4.6 Accuracy assessment of different species distribution models at the Penglaneinon site. AUC = Area under the curve, COR = Pearson's correlation, TSS = True skill statistics, Deviance = Deviance of the predicted values from the observed values. (±values) = difference in AUC for 10m scores relative to 0.5 m AUC scores. For clarity, both MLP and RBF are different ANN methods.

Resolution	Model	AUC	COR	TSS	Deviance
0.5m	GAM	0.74	0.40	0.37	1.13
	MARS	0.74	0.36	0.37	1.13
	SVM	0.72	0.36	0.36	1.56
	RF	0.82	0.53	0.50	0.95
	BRT	0.75	0.41	0.39	1.14
	MaxEnt	0.74	0.39	0.38	1.25
	MLP	0.73	0.38	0.37	1.38
	RBF	0.72	0.36	0.34	1.17
10m	GAM	0.78 (+0.04)	0.46	0.44	1.07
	MARS	0.77 (-0.03)	0.45	0.43	1.08
	SVM	0.77 (+0.05)	0.46	0.45	1.44
	RF	0.83 (+0.01)	0.56	0.54	0.93
	BRT	0.78 (+0.03)	0.47	0.44	1.11
	MaxEnt	0.78 (+0.04)	0.45	0.45	1.17
	MLP	0.77 (+0.04)	0.45	0.44	1.28
	RBF	0.76 (+0.04)	0.41	0.42	1.15

The predictive accuracy assessment for Penglaneinon proved similar to that of Ffridd Fawr in many ways. Whilst the average AUC scores were marginally lower (-0.018 AUC) at Ffridd Fawr (M = 0.745 AUC) then Penglaneinon (M = 0.763 AUC), all models were still classed as 'moderate' in performance (0.7-0.9 AUC). Like Ffridd Fawr, the difference in predictive accuracy between resolutions was observed to be similar with slight increases in performance (+0.035 AUC) for 10 m models (M= 0.78 AUC) against 0.5 m models (M = 0.745 AUC). RF again was estimated to be the most accurate across both resolutions, and RBF the least. Finally, there did not appear to be any notable differences in predictive accuracy between the different modelling method types (i.e. classical regression (M = 0.740 AUC) vs machine learning (M =0.746 AUC)).

4.5.2 Variable importance at different resolutions

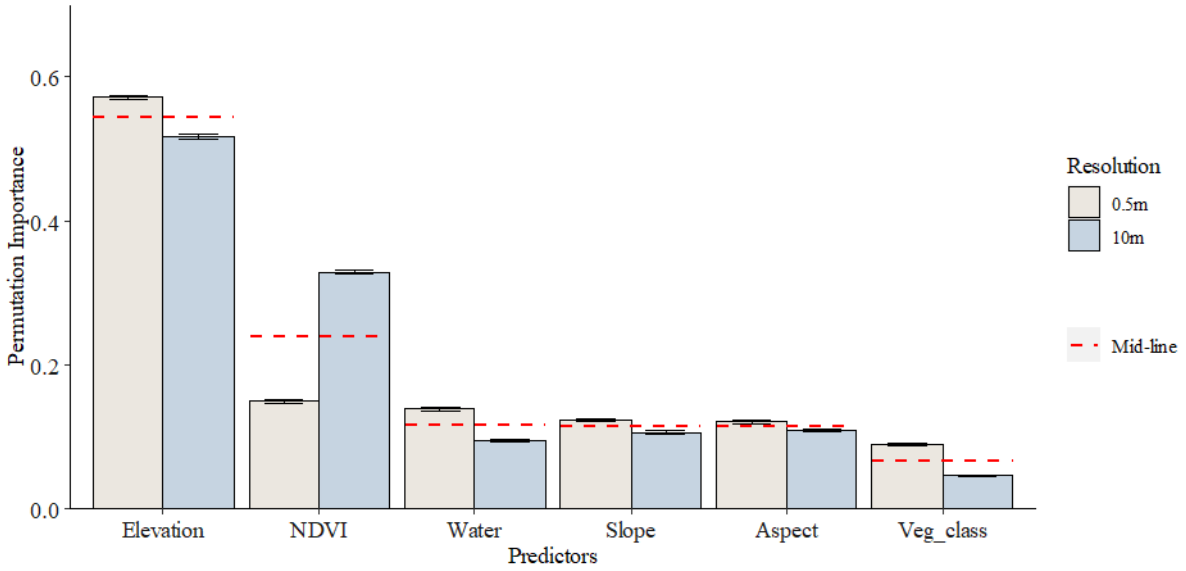


Figure 4.11 Mean variable importance across the 8 different models using different predictor resolutions at Penglaneinon. Predictors are ordered on the x axis according to the 0.5 m importance rank to facilitate comparison. Error bars = standard error of the mean (SEM).

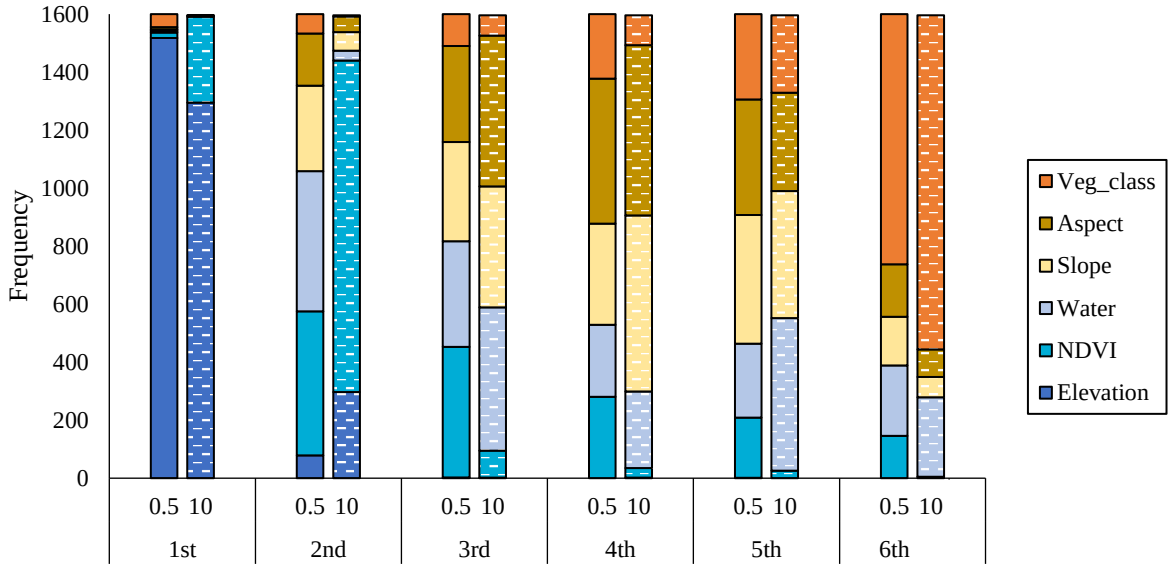


Figure 4.12 Frequency of occurrence for predictors at each level of the variable importance order according to resolution at Penglaneinon. Total number of model iterations for each resolution = 1600. Solid bars denote variables at 0.5 m resolution. Hatched bars denote variables at 10m resolution.

At Ffridd Fawr, three key predictors were identified ('water', 'slope', 'aspect') which were likely be having a notable effect on distribution. At Penglaneinon however, 'elevation' is clearly demonstrated to be the primary predictor affecting distribution. As demonstrated by Figure 4.12, 'elevation' was ranked first 95% (n = 1518) of the time using 0.5 m resolution predictors, and 80% using 10 m predictors (n = 1295). Furthermore, 'elevation' did not feature any lower than 2nd in importance

rankings at either resolution. It is evident from Figure 4.11 that there was consensus in variable importance for all predictors with the exception of NDVI. Whilst NDVI was ranked 2nd in 71% of model iterations (n =1143) using 10 m resolution predictors, it only accounted for 31% of iterations with 0.5 m resolution predictors (Figure 4.12). 'Aspect', 'slope' and 'water' feature interchangeably in 3rd, 4th and 5th rankings (Figure 4.12), whilst 'vegetation classification' appeared to be the least important (6th) variable affecting distribution (54% of model iterations (n = 862) at 0.5 m predictor resolution, 72% at 10 m predictor resolution (n =1152)).

The level of individual model variability in predictor rankings for Penglaneinon (Figure 4.13 & Figure 4.14) was generally less than it was for Ffridd Fawr, particularly at the 0.5 m predictor resolution. MLP outputs again are consistently higher than other models across all predictors at both resolutions, whilst SVM predicts notably lower permutation importance than other models for 'elevation'. Though permutation importance increases to a degree in all models for NDVI between 0.5 m and 10 m predictors, it is evident that much of the difference in average variable importance is due to the pronounced increase in SVM and BRT models.

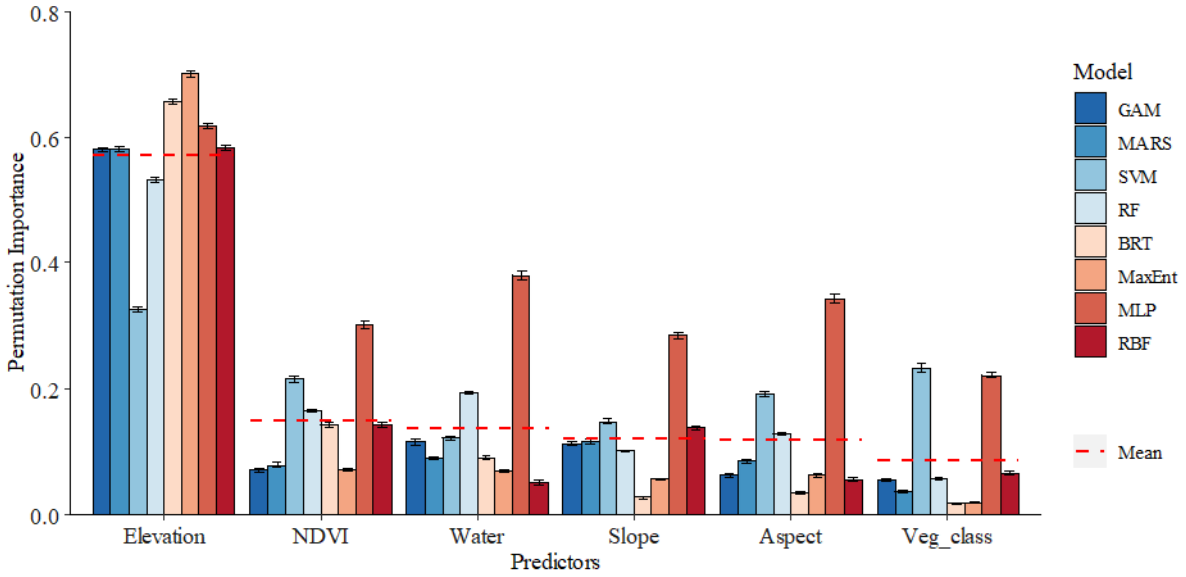


Figure 4.13 Mean variable importance for each predictor according to individual model types at 0.5 m predictor resolution at Penglaneinon. Each model is averaged across 200 iterations. Mean line = average across all models. Predictor order is arranged highest to lowest according to averaged variable importance.

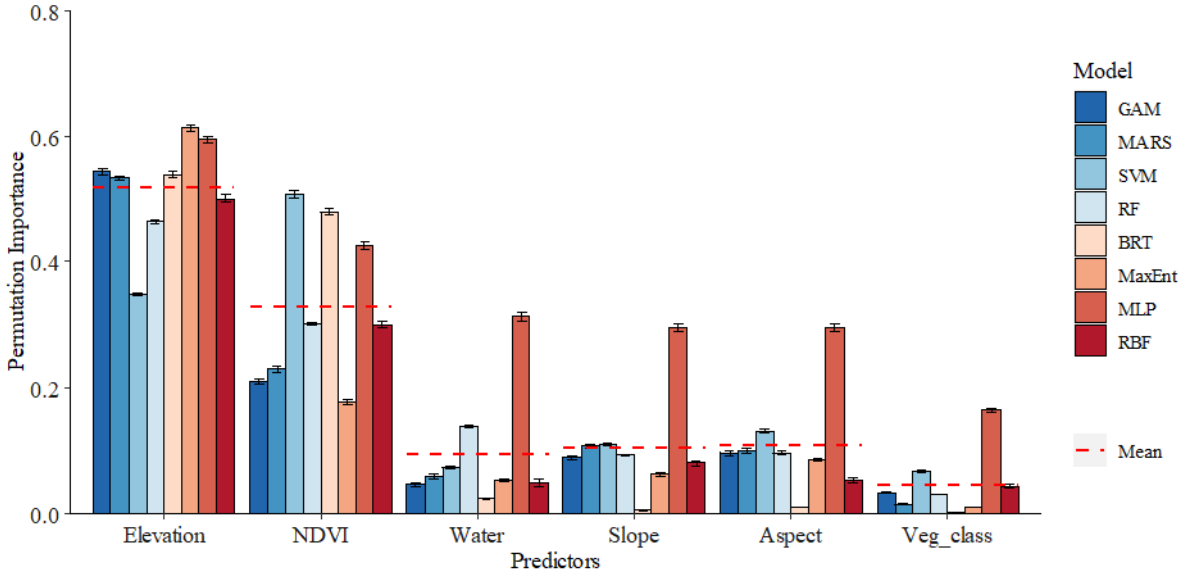


Figure 4.14 Mean variable importance for each predictor according to individual model types at 10 m predictor resolution at Penglaneinon. Each model is averaged across 200 iterations. Mean line = average across all models. Predictor order is arranged highest to lowest according to averaged variable importance of the 0.5 m resolution models to facilitate clear comparison.

Table 4.7 Predictor importance ranking according to each individual model type at both predictor resolutions. For clarity, both MLP and RBF are different ANN methods.

Resolution	Models	Predictors (ordered according to average variable importance)					
		Elevation (1st)	**NDVI (2nd)**	**Water (3rd)**	**Slope (4th)**	**Aspect (5th)**	**Veg_class (6th)**
0.5	GAM	1	4	2	3	5	6
	MARS	1	5	3	2	4	6
	SVM	1	3	6	5	4	2
	RF	1	3	2	5	4	6
	BRT	1	2	3	5	4	6
	MaxEnt	1	2	3	5	4	6
	MLP	1	4	2	5	3	6
	RBF	1	2	6	3	5	4
Resolution	**Models**	**Predictors (ordered according to average variable importance)**					
		Elevation (1st)	**NDVI (2nd)**	**Aspect (3rd)**	**Slope (4th)**	**Water (5th)**	**Veg_class (6th)**
10	GAM	1	2	3	4	5	6
	MARS	1	2	4	3	5	6
	SVM	2	1	3	4	5	6
	RF	1	2	4	5	3	6
	BRT	1	2	4	5	3	6
	MaxEnt	1	2	3	4	5	6
	MLP	1	2	5	4	3	6
	RBF	1	2	4	3	5	6

Table 4.7 reinforces the conclusion that ‘elevation’ was the most important variable affecting distribution-as all bar the 10 m resolution SVM model placed ‘elevation’ first in the rankings. Likewise, all bar two models (SVM and RBF at 0.5 m resolution) placed vegetation class last. Finally, whilst Figure 4.14 indicated large increases in permutation importance to be central to the overall mean increase in NDVI between 0.5m and 10m resolutions, it is notable from Table 4.7 that all models at a 10 m resolution placed NDVI second in importance (aside from SVM which places it first), thereby suggesting a broader consensus between models towards NDVI importance then just large increases in SVM and BRT models at the 10m resolution. However, the degree of variability in NDVI rankings at a 0.5m resolution is also prevalent, which suggests that the true overall importance

of NDVI cannot be discerned beyond input predictor resolution. Finally, it is clear at both resolutions that ‘aspect’, ‘slope’ and ‘water’ could not be differentiated in ranked importance from each other.

4.5.3 Response curves

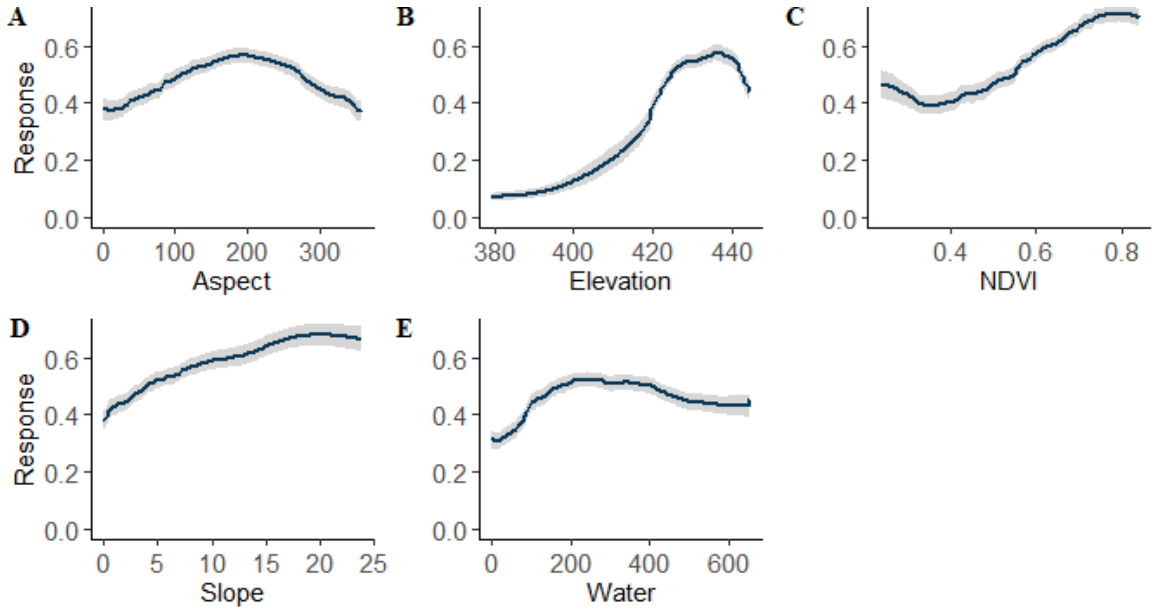


Figure 4.15 Mean response curves for all model types for each predictor at Penglaneinon using 0.5 m resolution predictors. Shaded areas denote standard error of the mean Vegetation classification was not included due to ‘sdm’ R library errors in prediction.

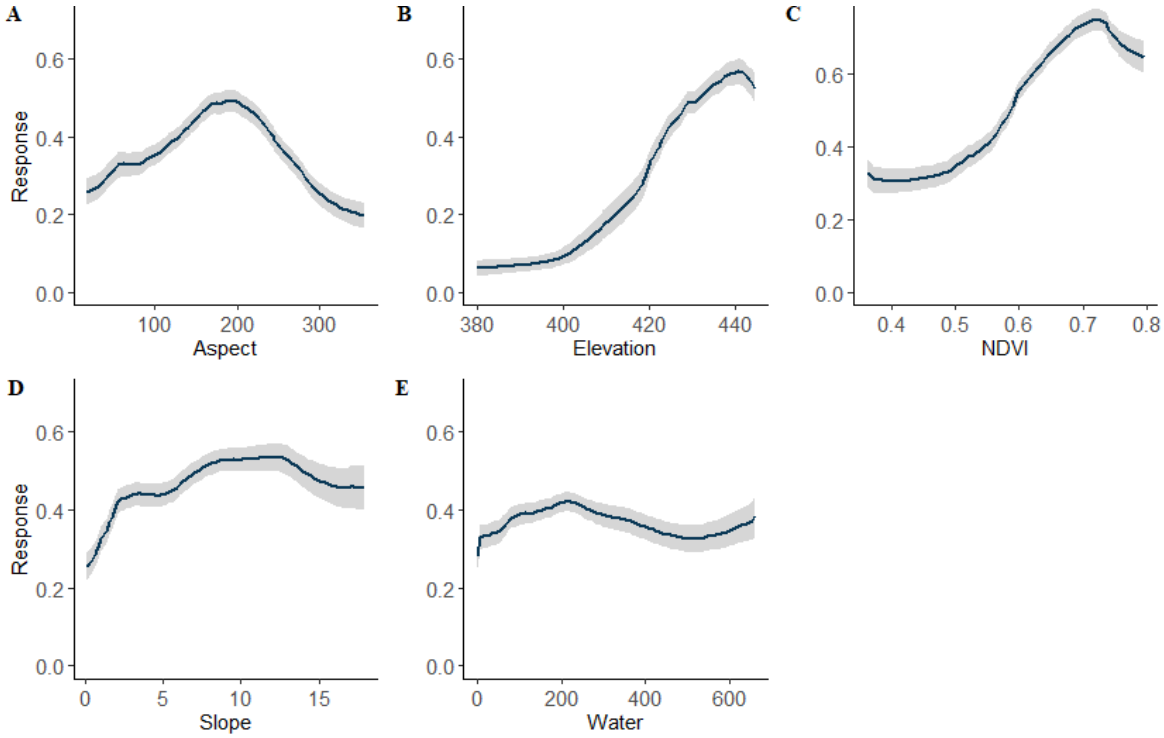


Figure 4.16 Mean response curves for all model types for each predictor at Ffridd Fawr using 10 m resolution predictors. NDVI was adjusted to 0.75 ymax (all others at 0.7) to capture the whole response curve. Vegetation classification was not included due to R ‘sdm’ library errors in prediction. Shaded areas denote standard error of the mean.

As was the case with the response curves at Ffridd Fawr, general relationships were consistent across the two resolutions, though some response curve shapes were complex forms (e.g. neither wholly linear nor unimodal). At an individual level, 'aspect' matched the hypothesised response of a unimodal shape, with probability of occurrence being more likely the nearer to south-facing (180°) slopes the animals were. The observed response for 'elevation' also concurred with the hypothesised response, as the probability of occurrence increased as elevation increased to ~440 m. It was confirmed by visual inspection in QGIS that the decrease in response witnessed at the highest altitudes related to small areas which appeared to be outcrops. However, on-site confirmation would be required to verify this. In permutation importance results, NDVI was most dissimilar between resolutions. From analysing the response curves at both resolutions, the response follows a rough s-shaped response. Despite the complex form, the general relationship (response increases as NDVI values increases) matched the hypothesised response. The differences between resolutions were notable in the initial decrease at the 0.5 m resolution to NDVI = ~0.35 which did not happen at the 10 m resolution, whilst the level of response was also larger at the 10 m (~0.35 - 0.75) than at the 0.5 m (~0.45 - 0.7) resolution. Though further investigation is required, it is likely that this uncertainty between resolutions was being caused by the spectral properties of the *Molinia*. At the time of image surveying, the new *Molinia* growth was still establishing, and therefore areas of the swards were interspersed with both fresh green *Molinia* and yellow senesced material bult up across previous years. At a 0.5 m resolution, this variation in sward was apparent, however at a 10 m level much of this variation had been lost. As such, the response of the animals to variations in *Molinia* growth stages between different resolution is much more uncertain.

The response to 'slope' was the most varied of all the predictors between the two different resolutions. At the 0.5 m level, a small linear increase in response was predicted as slope increased, whereas the 10 m response appeared to correspond to a more complex form that was part unimodal and linear in shape. Furthermore, neither form matched the hypothesised response of linear decrease as slope increased, but as the response was small in both cases, it concurs with the variable importance ranking of 'slope' being uninfluential in effecting distribution. Finally, the overall response for 'water' also appeared negligible. It was hypothesised that probability of occurrence (response) would decrease as distance from water increased. However, the only notable response that appears to occur is in the first ~200 m from water, where contrastingly the response increased as distance from water increased. From 200 m onwards, tentative observation suggests that response declines as distance to water increases, however the variation in form between resolutions creates uncertainty in this statement. From knowledge gained through site visits, plus further investigations using QGIS, it is postulated that the initial 200 m increase in response from water is due to much of the main stream being located in a slight gulley, with a riparian corridor of dense *Molinia* making much of the stream inaccessible. It is reasoned that the animals are using select fords (crossings) to

access water (when not near to the blanket bog), which in part is causing a low density of locations to be observed near much of the stream (Figure 4.17). The exact reason for the lack of clear response beyond 200 m distance is unclear without further investigation. A primary theory is that the degree of access to water on site is not restrictive enough for animal distribution to require being closer to water bodies. Though the vegetation density at Penglaneinon would likely restrict movement considerably, an accompanying MSc study (supervised on a day-to-day basis by the PhD candidate) revealed the site to contain a network of pathways created by the sheep which indicated that the animals may be able to navigate larger distances across the site efficiently (Noble 2018). It is therefore plausible that the animals could reside away from water, but still able to easily access it when required.

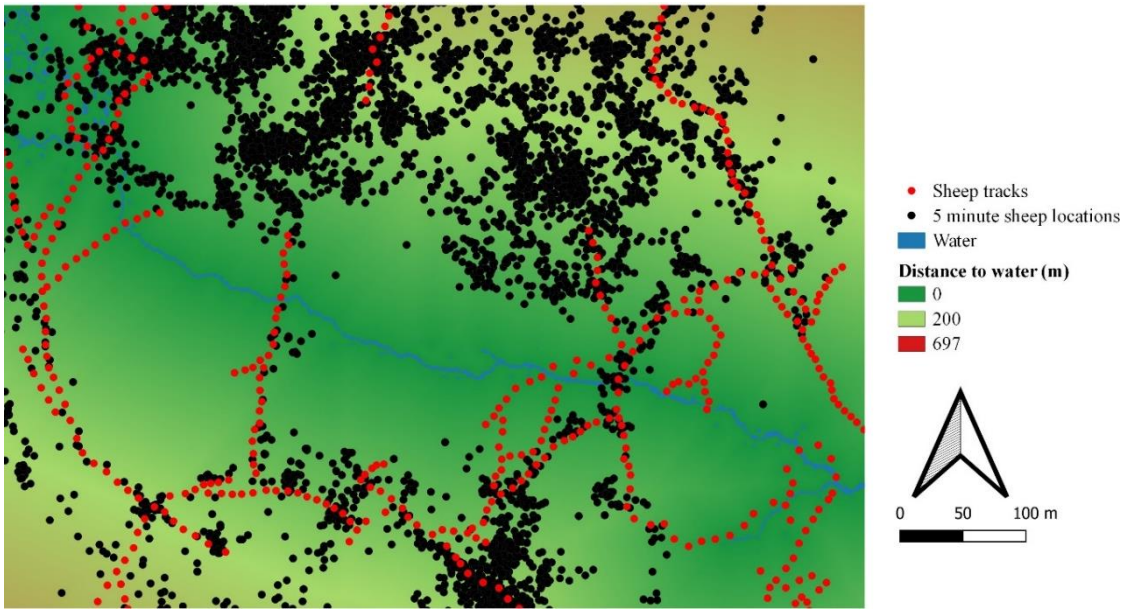


Figure 4.17 Stream crossings at Penglaneinon site indicated by mapped pathways and 5 min interval sheep location data.

4.6 Chapter outcomes summary

4.6.1 Summary of results and proposed future improvements for species distribution modelling at Ffridd Fawr and Penglaneinon

In summarising the findings for Ffridd Fawr, it is evident from the variable importance measurements and response curve forms that ‘water’, ‘standing shelter’ and ‘NDVI’ could be considered key factors affecting distribution of animals at this site at the time of study. Furthermore, there is clear indication that ‘slope’, ‘aspect’ and ‘topographic shelter’ are not having a meaningful impact on distribution. Though only a moderate response to ‘elevation’ was detected, the nature of the response curve would suggest potential unexplained influence. It is possible that as all three key factors feature almost exclusively in the northern side of the site, that the response to ‘elevation’ is incidental, however further investigations would be required to verify this. Studying climate (e.g. temperature, wind, rain) directly across the site would present a key route of exploration. Linked to this could be the creation of ‘cost’ rasters. These would seek to measure the physiological cost of traversing a given area of the site and could be integrated with least-cost path analysis. Such a map would look to account for topography (slope and elevation), climate, vegetation thickness, and obstructions which alter movement (e.g. streams, ditches). Finally, conducting a vegetation classification as originally planned in this study would also be an obvious inclusion.

At Penglaneinon, it is evident that only ‘elevation’ can be considered a key factor affecting distribution. Though the response curves suggest NDVI to have a notable effect, the variation between predictor resolutions means that the SDM outputs alone cannot be certain. A next key step in future research would be to explore what proximal factors elevation is likely to be representing at this site. As suggested for Ffridd Fawr, measuring climatic variables directly would be a major improvement. A second option would be to explore visibility on the site and attempt to assess whether animals are seeking higher elevations to better detect potential predators (e.g. especially as the ewes had lambs), and/or resources. If vegetation height could be estimated, it would be possible to combine this with topology and perform viewshed type analysis to determine at sheep height what visibility was like at different locations. As suggested for Ffridd Fawr, a cost raster would also be beneficial for Penglaneinon, especially as the vegetation is very dense for much of the site, and the role of the path network present is still not fully understood. Deriving the cost raster and then calculating least-cost paths along the site would facilitate a comparison with the known pathways to better understand their use and role. Given the use of UAV technologies already in the methodology, measuring vegetation height could be explored using methodologies such as those outlined in Forsmoo *et al* (2018). However, another suggestion given the density of the sward at Penglaneinon would be to selectively cut sections of the sward to define the terrain, then use either LIDAR or structure through motion

(SfM) processes in image stitching to derive a digital surface model (DSM), which could then be subtracted from a digital terrain model (DTM) to give above ground height.

4.6.2 Improvements to the modelling process

Though species distribution models with ‘moderate’ predictive ability are well witnessed in the literature (Svenning, Normand & Kageyama 2008; Byrne *et al.* 2014), improvements in the AUC scores for the SDM models at both sites could likely be achieved with further individual tuning of the models. In many cases default parameters were used, whereas iteratively exploring different settings for models (e.g. different number of trees in random forest models) could likely lead to more accurate models (Hao *et al.* 2020).

Another shortcoming would be that the models were run for the whole tracking datasets for each site. Given the aim of the exercise is to be able to manipulate animal behaviour, a shift towards a dynamic and even mechanistic modelling in future would be recommended. Given the seasonal variation in climatic conditions between years, it is possible that basing proposed manipulations on data from previous years would lead to less effective outcomes. A more reasoned approach would be to collect data and perform SDMs at periodic intervals within a season and alter interventions dynamically throughout the season. An added benefit of this method would be that each subsequent recording interval could act as independent data sets with which to validate the previous SDMs; adding additional rigour to the model assessment beyond internal validation (Franklin 2010a). Though this type of modelling requires far more physiological (Kearney & Porter 2009) (e.g. diet, body temperature) and ecological information (Guisan & Zimmermann 2000; Franklin 2010b) to implement successfully, the small scale of the sites (compared to most other SDM studies), and known history of the animals being studied offers considerable advantages to achieving this.

Finally, the resolution by which behaviour should be studied needs to be explored. Interpretation of the results in this study considered two input resolutions so positional uncertainty could be accounted for. However, the level of variation in predictor importance between resolutions prevented detailed post-hoc analysis (e.g. pairwise-comparisons) on specific interactions between predictors being assessed. Though differential correction of GNSS locations (most likely through post-processed kinematic (PPK) workflow) could lessen the likely positional uncertainty in the data, the computational expense of processing such high resolution predictors within an SDM means that high granularity should not merely be used for the sake of it. Resolution should instead be determined by the study objectives and site scale, but also importantly the ecological scale that best matches the study animal. For example, at what spatial resolution does grazing become meaningful for sheep? This is not something well discussed in the literature as the majority of studies have utilised much lower resolution data. Manzoor, Griffiths & Lukac (2018) note that most studies choose grain size based on the data available, rather than relevant factors such as species ecology or spatial scale of

study. This is partly because choice of predictors is ‘frequently opportunistic’, and reliant on a standard combination of 19 bioclimatic variables (Porfirio *et al.* 2014). Given these are all a minimum 1 km GSD resolution, this means that all other biophysical variables (e.g. NDVI) must be resampled to this size (Manzoor, Griffiths & Lukac 2018). As such, to date very few studies have utilised fine scale predictors (<10 m resolution) in SDMs (Descombes *et al.* 2016), and therefore the discussion as to what resolution is actually prudent for interpretation is underdeveloped.

CHAPTER 5

GENERAL DISCUSSION

5.1 A summary of the aims and objectives of the PhD study

This study aimed to develop a standardised workflow (encompassed in the objectives set out in Table 1.2) for collecting and analysing data that would enable site specific habitat restoration techniques to be recommended for larger-scale (+50 ha) sites. In particular, the study focused on measuring the relationship between grazing animals and the environmental factors driving their distribution. The theory was that targeted grazing prescriptions could be better implemented if practitioners knew which factors to manipulate. It was reasoned that this could be an improved approach over pre-existing prescription systems (such as LSUs) which make assumptions about the uniformity of grazing over a site. All the individual objectives for each chapter (specified in Chapter 1) were completed, and as such the three main study aims were successfully addressed. Each site was comprehensively profiled, with likely factors affecting sheep distribution baselined. Animals were successfully tracked on both sites using low-cost GNSS collars, with a new UAVRTS developed for whole herd monitoring. Data processing and behavioural inference tools were also developed, whilst grazing intensity was analysed using a combination of home range analyses, and a bespoke true LSU indicator across the site. Finally, environmental factors were assessed with regards to their effect on the spatial distribution of the animals, with a statistical evaluation carried out and the ecological realism of the results considered.

5.2 Recommending management interventions based on the workflow

Fundamentally, the true novelty of the developed workflow is the level of insight offered on the effect environment variables are having on animal distribution on large sites. The amount of data produced, and the ability to utilise different combinations of newly developed/existing tools (e.g. behavioural inference, species distribution modelling, home range estimation) to understand how distribution and grazing behaviour is being driven on a specific site, lends itself as an effective tool for recommending targeted management interventions. The level of interpretation possible is not limiting, with the recommendations for interventions ultimately dependent on the specific objectives for a site (i.e. which ecosystem services are being targeted). For example, in this study it was clear from home range and LU estimation that at both sites grazing intensity was not uniform. This does not necessarily infer grazing was having a detrimental effect, as it depends on the precise conservation objectives set for the site. At Ffridd Fawr, the desire to manage the sward to consist of areas of both short and long vegetation for *Numenius arquata* (curlew), means non uniform grazing may well be preferential. However, at Penglaneinon, the objective is to reduce the overall dominance of *Molinia* on the site. Visual observations at Penglaneinon on QGIS suggested that animals were focusing on previously cut areas, the result of this was that large areas of monoculture *Molinia* were barely impacted. Manipulation of the grazing behaviour would therefore be more beneficial on this site to provide more even grazing pressure and prevent the animals from over-impacting vulnerable areas

(e.g. the blanket bog). Results from the species distribution modelling indicated NDVI to have a notable effect on distribution. Combining this information with the insight gleaned from the observed localised grazing intensity for the site in Chapter 3, would support the notion of further cutting as a way of achieving this. This example demonstrates the utility of adopting multiple analyses for interpretation, and also reinforces the notion that management will ultimately be most effective when clear management objectives are aligned with comprehensive data driven interpretation (as provided with this workflow).

5.3 Limitations of the workflow and areas of future research

Though the creation of the workflow was ultimately successful, the time and cost of the hardware/software development undertaken effectively curtailed the possibility of imposing any interventions. Further research should ultimately aim to conduct interventions after initial baselining, as this would provide critical ground truth evidence as to the effectiveness of any recommendations derived from the workflow.

Aside from imposing interventions, another key area of development should be to introduce a more dynamic component into the analyses. Within this study, data was collected during the growing season, then subsequently analysed. However, the limitation of this is that any proposed interventions would not be implemented until the following year, as in many cases livestock are moved from upland sites during winter (i.e. the sites in this study). Given the differences in climatic conditions between seasons from one year to the next, drivers of animal distribution may well alter. Therefore, prescribing interventions on the basis of data from a previous year may well prove ineffectual. As noted in Chapter 4, a key flaw of species distribution modelling is that they assume equilibrium (Guisan & Zimmermann 2000; Austin 2002), and their static nature has often being criticised for not representing the constant dynamic nature of animal-environment interactions (Guisan & Theurillat 2000; Guisan & Thuiller 2005). As such, a key factor in progressing this approach would be to collect and analyse data at regular intervals within a season (Fernandez *et al.* 2017), and impose interventions in response to each collection episode. By doing this, the effects of seasonality would be negated (e.g. different stages of vegetation growth), and the extent to which factors affecting distribution dynamically alter in their importance during a season could be explored (Zuckerberg *et al.* 2016).

There may be a concern that repeated data collection over a season could prove laborious, however future improvements in the technology used could counteract this. The use of UAVs already means that repeated surveys within a season to measure environmental variables are already possible. With regards to the tracking operations, the use of the UAVRTS would be particularly useful to collect a snapshot of the behaviour of all the animals periodically at the same time as the image surveys are being conducted. In addition, further research into developing the UAVs into data mules (mobile data

collectors) has already received funding (B. Roberts via the Joy Welch Trust). This will enable data to be uploaded from the GNSS tracking collars (and micro-climate sensors if employed) directly on to a UAV when passing overhead, and as such minimise the amount of contact required with the animals. This is particularly useful on large sites, where gathering in collared animals can take a considerable amount of time. The other benefit to utilising this approach is that all data required (i.e. image surveying and tracking data) would effectively be collected within each single survey.

Additional improvements to the workflow could include further improvement of the Beferrer/Shedsplitter software to enhance user friendliness. Use of the software for data in later studies has highlighted the inflexibility of the software to different data formats and structures, and also that the command line interface could be improved upon. As such, further work beyond this study has already begun on creating a Graphical User Interface (GUI) which should promote accessibility to the software as well as allowing different data inputs to be defined.

Aside from the technological limitations and improvements to the methodologies used within workflow, there would be a number of opportunities for other disciplines to be used to compliment the approach. One area for development mentioned in Chapter 4 would be adopting a more mechanistic approach to the species distribution modelling (Kearney & Porter 2009). The current workflow only assesses the relationship between species distribution and environmental variables. It does not assess how the physiology of the animal alters to changes in environment. This knowledge would enable a more complete understanding of animal-environment interactions, as the wider effect of environmental variables on an animal would be better understood, and not just the effect on movement. Understanding of diet selection under different environmental conditions would be a key aspect to consider (Peterson, Papeş & Soberón 2015). This could be achieved using methodologies such as faecal n-alkane analysis (Fraser *et al.* 2009), faecal cuticle analysis (Stevens *et al.* 1987) or DNA barcoding of faecal matter (Pegard *et al.* 2009; Camp *et al.* 2020) to identify and quantify the species of plants being consumed. Complimenting this could be measurements of body temperature (e.g. using a thermal camera mounted on a UAV (Kays *et al.* 2019)) and heart rate monitoring (e.g. using sensors attached to the body (Andersson 2016)). Understanding how body temperature and heart rate alters (alongside weight and condition measurements at the beginning and end of study) could indicate the level of physiological stress the animals were experiencing at any one time (Al-Haidary 2004; Maurya *et al.* 2007). Combining this information with the knowledge of the environmental conditions would therefore facilitate a far more complete understanding as to why grazing behaviour changes with changes in environment. This would not only be beneficial for habitat management, but also more widely for animal husbandry and agriculture.

As the workflow has been developed to offer site specific information, there is also a question of scalability, particularly in upland areas. Given that peatland/heather moorland areas often encompass

a number of large enclosures spanning thousands of hectares (as is the case for Penglaneinon at the Elan valley, and Ffridd Fawr within the Lake Vyrnwy estate), it may not be feasible to utilise UAVs across all areas. As such it may be necessary to adopt different surveying tools that are capable of covering such large areas. One such recommendation would be use UAV data to build a comprehensive understanding of the main environmental drivers at a single enclosure, then utilise high resolution satellite imagery on the adjacent enclosures to scale up. The enclosure surveyed by UAV data could act as the reference site, with the knowledge gained applied over the remaining enclosures with the satellite imagery used to monitor impacts.

5.4 The need to know the manipulation capacity of a site

A final observation is that it is important to recognise the ability of a site to reach the targeted management objectives (Ehrenfeld 2000). Understanding the extent to which a site can be manipulated helps manage expectations of what any interventions are likely to achieve (Lake 2001). For example, whilst the objective at Penglaneinon was to reduce the *Molinia* dominance so that peatland vegetation could return, it may well be that the natural composition of the site is upland fringe, with multiple habitat types present. Further expansion of the workflow could therefore focus on developing methodologies to assess the manipulation capacity of a site. Examples could include locating historical data to inform on previous site composition, and the use of ecological niche/bioclimatic modelling of target species (e.g. *Sphagnum* moss for peatlands (Gignac, Halsey & Vitt 2000)) to estimate the likely suitable habitat of the species on the site under current environmental conditions. The advantage of this particular approach is that modelling could be extended to include suitability under future climates, which would prove considerably valuable given that habitats such as peatlands often require a considerable amount of time to return to proper working function, and therefore any future changes in climate on a site should be accounted for (Oke & Hager 2017). With this knowledge, interventions can then focus on altering a site within the realistic limits, and not beyond these. The success of any management intervention can then be assessed by the capacity a site has for manipulation, and not simply by the target objective set upon it. As Hobbs *et al* (2011) note,

"*Scientists who wish to alter the function of ecosystems to enhance the services they deliver (including, especially, maintenance of a great diversity of organisms) should be explicit about both their goals and the possibilities of achieving them. We should not, for instance, give the impression that ways can be found to either hold or turn back the clock and preserve or recreate imagined Edens*"

5.5 Conclusions

Though future research is required to comprehensively ground truth the workflow, the amount of data produced and subsequent insight attainable demonstrates the true novelty of the approach, particularly given the low input cost. This ultimately affirms the overall central research question posed in the PhD study (*'Can recent developments in sensing technologies and/or processing capability be harnessed to improve the evidence-base for conservation decision-making?'*).

By estimating the characteristics of grazing behaviour and the specific environmental variables driving distribution on a site, the workflow developed offers practitioners a greater understanding of animal-environment interactions over large-scale sites. Though upland habitats and peatlands in particular were the focus for this work, the nature of the approach enables it to be used in a multitude of different site/habitat scenarios. As such the workflow represents a potential option for many varied land management scenarios globally.

REFERENCES

Acs, S., Hanley, N., Dallimer, M., Gaston, K.J., Robertson, P., Wilson, P. & Armsworth, P.R. (2010) The effect of decoupling on marginal agricultural systems: implications for farm incomes, land use and upland ecology. *Land Use Policy,* **27,** 550-563.

Adler, P., Raff, D. & Lauenroth, W. (2001) The effect of grazing on the spatial heterogeneity of vegetation. *Oecologia,* **128,** 465-479.

Al-Haidary, A.A. (2004) Physiological responses of Naimey sheep to heat stress challenge under semi-arid environments. *International Journal of Agriculture and Biology,* **2,** 307-309.

Andersen, R., Farrell, C., Graf, M., Muller, F., Calvar, E., Frankard, P., Caporn, S. & Anderson, P. (2017) An overview of the progress and challenges of peatland restoration in Western Europe. *Restoration Ecology,* **25,** 271-282.

Anderson, D.M., Estell, R.E. & Cibils, A.F. (2013) Spatiotemporal cattle data—a plea for protocol standardization. *Positioning,* **4,** 115.

Andersson, I. (2016) Behaviour and heart rate in sheep when herded by Border collies with different background. Second cycle, A2E. Skara: SLU, Dept. of Animal Environment and Health.

Animut, G. & Goetsch, A. (2008) Co-grazing of sheep and goats: benefits and constraints. *Small Ruminant Research,* **77,** 127-145.

Apitz, S.E. (2013) Ecosystem services and environmental decision making: Seeking order in complexity. *Integrated Environmental Assessment and Management,* **9,** 214-230.

Araújo, M.B. & New, M. (2007) Ensemble forecasting of species distributions. *Trends in Ecology & Evolution,* **22,** 42-47.

Araújo, M.B. & Peterson, A.T. (2012) Uses and misuses of bioclimatic envelope modeling. *Ecology,* **93,** 1527-1539.

Armstrong, A., Holden, J., Kay, P., Foulger, M., Gledhill, S., McDonald, A. & Walker, A. (2009) Drain-blocking techniques on blanket peat: A framework for best practice. *Journal of Environmental Management,* **90,** 3512-3519.

Asner, G.P., Elmore, A.J., Olander, L.P., Martin, R.E. & Harris, A.T. (2004) Grazing systems, ecosystem responses, and global change. *Annual Review of Environment and Resources,* **29,** 261-299.

Atkins, J.P., Burdon, D., Elliott, M. & Gregory, A.J. (2011) Management of the marine environment: integrating ecosystem services and societal benefits with the DPSIR framework in a systems approach. *Marine Pollution Bulletin,* **62,** 215-226.

Austin, M. (2002) Spatial prediction of species distribution: an interface between ecological theory and statistical modelling. *Ecological Modelling,* **157,** 101-118.

Austin, M. (2007) Species distribution models and ecological theory: a critical assessment and some possible new approaches. *Ecological Modelling,* **200,** 1-19.

Averis, A. (2004) *Illustrated Guide To British Upland Vegetation*. Joint Nature Conservation Committee.

Avgar, T., Potts, J.R., Lewis, M.A. & Boyce, M.S. (2016) Integrated step selection analysis: bridging the gap between resource selection and animal movement. *Methods in Ecology and Evolution,* **7,** 619-630.

Bain, C., Bonn, A., Stoneman, R., Chapman, S., Coupar, A., Evans, M., Gearey, B., Howat, M., Joosten, H. & Keenleyside, C. (2011) *IUCN UK commission of inquiry on peatlands*. IUCN UK Peatland Programme.

Bakker, J.P. (1989) *Nature management by grazing and cutting*. Kluwer Academic Publishers.

Bangs, P.D., Krausman, P.R., Kunkel, K.E. & Parsons, Z.D. (2005) Habitat use by desert bighorn sheep during lambing. *European Journal of Wildlife Research,* **51,** 178-184.

Bayram, H., Stefas, N. & Isler, V. (2018) Aerial radio-based telemetry for tracking wildlife. *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 4723-4728. IEEE.

Bean, W.T., Stafford, R. & Brashares, J.S. (2012) The effects of small sample size and sample bias on threshold selection and accuracy assessment of species distribution models. *Ecography,* **35,** 250-258.

Beerling, D.J., Huntley, B. & Bailey, J.P. (1995) Climate and the distribution of *Fallopia japonica*: use of an introduced species to test the predictive capacity of response surfaces. *Journal of Vegetation Science,* **6,** 269-282.

Beyer, F., Jurasinski, G., Couwenberg, J. & Grenzdörffer, G. (2019) Multisensor data to derive peatland vegetation communities using a fixed-wing unmanned aerial vehicle. *International Journal of Remote Sensing,* **40,** 9103-9125.

Bhardwaj, A., Sam, L., Martín-Torres, F.J. & Kumar, R. (2016) UAVs as remote sensing platform in glaciology: Present applications and future prospects. *Remote Sensing Of Environment,* **175,** 196-204.

Bird, P., Lynch, J. & Obst, J. (1984) Effect of shelter on plant and animal production. *Animal Production in Australia,* **15,** 270-273.

Bivand, R. & Rundel, C. (2017) rgeos: interface to geometry engine-open source (GEOS). *R package, Version 0.3-23*.

Bonn, A., Holden, J., Parnell, M., Worrall, F., Chapman, P., Evans, C., Termansen, M., Beharry-Borg, N., Acreman, M. & Rowe, E. (2010) *Ecosystem services of peat - Phase 1*. Defra, 140pp. (CEH Project Number: C03758 T1)

Boria, R.A., Olson, L.E., Goodman, S.M. & Anderson, R.P. (2014) Spatial filtering to reduce sampling bias can improve the performance of ecological niche models. *Ecological Modelling,* **275,** 73-77.

Bourgeau-Chavez, L.L., Endres, S., Powell, R., Battaglia, M.J., Benscoter, B., Turetsky, M., Kasischke, E.S. & Banda, E. (2017) Mapping boreal peatland ecosystem types from multitemporal radar and optical satellite imagery. *Canadian Journal of Forest Research,* **47,** 545-559.

Boyce, M.S. (2006) Scale for resource selection functions. *Diversity and Distributions,* **12,** 269-276.

Boyce, M.S., Pitt, J., Northrup, J.M., Morehouse, A.T., Knopff, K.H., Cristescu, B. & Stenhouse, G.B. (2010) Temporal autocorrelation functions for movement rates from global positioning system radiotelemetry data. *Philosophical Transactions of the Royal Society B: Biological Sciences,* **365,** 2213-2219.

Bro-Jørgensen, J., Brown, M.E. & Pettorelli, N. (2008) Using the satellite-derived normalised difference vegetation index (NDVI) to explain ranging patterns in a lek-breeding antelope: the importance of scale. *Oecologia,* **158,** 177-182.

Brown, J.L. (2014) SDM toolbox: a python-based GIS toolkit for landscape genetic, biogeographic and species distribution model analyses. *Methods in Ecology and Evolution,* **5,** 694-700.

Buchanan, G.M., Pearce-Higgins, J.W., Douglas, D.J. & Grant, M.C. (2017) Quantifying the importance of multi-scale management and environmental variables on moorland bird abundance. *Ibis,* **159,** 744-756.

Buisson, L., Thuiller, W., Casajus, N., Lek, S. & Grenouillet, G. (2010) Uncertainty in ensemble forecasting of species distribution. *Global Change Biology,* **16,** 1145-1157.

Bunce, R. (1987) The extent and composition of upland areas in Great Britain. In: Bell, M.; Bunce, R. G. H.,(eds.) Agriculture and conservation in the hills and uplands. Grange-over-Sands, NERC/ITE, 19-21. (ITE Symposium, 23).

Bunce, R., Barr, C., Clarke, R., Howard, D. & Scott, A. (2007) ITE land classification of Great Britain 2007. NERC-Environmental Information Data Centre 1 May 2007, http://doi.org/10.5285/5f0605e4-aa2a-48ab-b47c-bf5510823e8f [Output (Electronic)]

Bunce, R.G., Wood, C.M. & Smart, S.M. (2018) The Ecology of British Upland landscapes. I. Composition of landscapes, habitats, vegetation and species. *Journal of Landscape Ecology,* **11,** 120-139.

Bunting, P., Clewley, D., Lucas, R.M. & Gillingham, S. (2014) The remote sensing and GIS software library (RSGISLib). *Computers & geosciences,* **62,** 216-226.

Burgman, M.A. & Fox, J.C. (2003) Bias in species range estimates from minimum convex polygons: implications for conservation and options for improved planning. *Animal Conservation forum*, pp. 19-28. Cambridge University Press.

Burt, W.H. (1943) Territoriality and home range concepts as applied to mammals. *Journal of Mammalogy,* **24,** 346-352.

Byrne, A.W., Acevedo, P., Green, S. & O'Keeffe, J. (2014) Estimating badger social-group abundance in the Republic of Ireland using cross-validated species distribution modelling. *Ecological Indicators,* **43,** 94-102.

Báldi, A. (2008) Habitat heterogeneity overrides the species–area relationship. *Journal of Biogeography,* **35,** 675-681.

Calabrese, J.M., Fleming, C.H. & Gurarie, E. (2016) ctmm: An R package for analyzing animal relocation data as a continuous-time stochastic process. *Methods in Ecology and Evolution,* **7,** 1124-1132.

Calenge, C. & Fortmann-Roe, S. (2014) adehabitatHR: Home Range Estimation. R package version 0.4. 11.

Camp, A., Croxford, A.E., Ford, C.S., Baumann, U., Clements, P.R., Hiendleder, S., Woolford, L., Netzel, G., Boardman, W.S. & Fletcher, M.T. (2020) Dual-locus DNA metabarcoding reveals southern hairy-nosed wombats (*Lasiorhinus latifrons Owen*) have a summer diet dominated by toxic invasive plants. *PloS one,* **15,** e0229390.

Chang, K.-t. & Tsai, B.-w. (1991) The effect of DEM resolution on slope and aspect mapping. *Cartography and geographic information systems,* **18,** 69-77.

Cliff, O.M., Fitch, R., Sukkarieh, S., Saunders, D.L. & Heinsohn, R. (2015) Online localization of radio-tagged wildlife with an autonomous aerial robot system. *Robotics: Science and Systems*.

Cohen, B.S., Prebyl, T.J., Collier, B.A. & Chamberlain, M.J. (2018) Home range estimator method and GPS sampling schedule affect habitat selection inferences for wild turkeys. *Wildlife Society Bulletin,* **42,** 150-159.

Cohen, W.B. & Goward, S.N. (2004) Landsat's role in ecological applications of remote sensing. *Bioscience,* **54,** 535-545.

Committee, J.N.C.C (2009) Common standards monitoring guidance for upland habitats. *DEFRA, London*.

Cooper, M.D., Evans, C.D., Zielinski, P., Levy, P.E., Gray, A., Peacock, M., Norris, D., Fenner, N. & Freeman, C. (2014) Infilled ditches are hotspots of landscape methane flux following peatland re-wetting. *Ecosystems,* **17,** 1227-1241.

Coudun, C. & Gégout, J.-C. (2006) The derivation of species response curves with Gaussian logistic regression is sensitive to sampling intensity and curve characteristics. *Ecological Modelling,* **199,** 164-175.

Crossman, N.D. & Bass, D.A. (2008) Application of common predictive habitat techniques for post-border weed risk management. *Diversity and Distributions,* **14,** 213-224.

DEFRA (2011) Uplands Policy Review. Policy paper (DEFRA ref. PB13456). Department for Environment, Food & Rural Affairs.

Descombes, P., Petitpierre, B., Morard, E., Berthoud, M., Guisan, A. & Vittoz, P. (2016) Monitoring and distribution modelling of invasive species along riverine habitats at very high resolution. *Biological Invasions,* **18,** 3665-3679.

Ditmer, M.A., Vincent, J.B., Werden, L.K., Tanner, J.C., Laske, T.G., Iaizzo, P.A., Garshelis, D.L. & Fieberg, J.R. (2015) Bears show a physiological but limited behavioral response to unmanned aerial vehicles. *Current Biology,* **25,** 2278-2283.

Dobos, R., Taylor, D., Trotter, M., McCorkell, B., Schneider, D. & Hinch, G. (2015) Characterising activities of free-ranging Merino ewes before, during and after lambing from GNSS data. *Small Ruminant Research,* **131,** 12-16.

Dormann, C.F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., Marquéz, J.R.G., Gruber, B., Lafourcade, B. & Leitão, P.J. (2013) Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography,* **36,** 27-46.

Dos Santos, G.A.M., Barnes, Z., Lo, E., Ritoper, B., Nishizaki, L., Tejeda, X., Ke, A., Lin, H., Schurgers, C. & Lin, A. (2014) Small unmanned aerial vehicle system for wildlife radio collar

tracking. *Mobile Ad Hoc and Sensor Systems (MASS), 2014 IEEE 11th International Conference on*, pp. 761-766. IEEE.
Dowle, M., Srinivasan, A., Short, T. & Lianoglou, S. (2017) data. table: Extension of data. frame. *R package version*, **1**.4 (2017)
Ehrenfeld, J.G. (2000) Defining the limits of restoration: the need for realistic goals. *Restoration Ecology*, **8,** 2-9.
Elith, J. & Leathwick, J.R. (2009) Species distribution models: ecological explanation and prediction across space and time. *Annual Review of Ecology, Evolution, and Systematics*, **40,** 677-697.
Elith, J., H. Graham, C., P. Anderson, R., Dudík, M., Ferrier, S., Guisan, A., J. Hijmans, R., Huettmann, F., R. Leathwick, J. & Lehmann, A. (2006) Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, **29,** 129-151.
Elton, C. (1927) Animal ecology In: Whittaker R, Levin S (eds) Niche: theory and application. Dowden, Hutchinson and Ross. *Inc, Stroudsburg, PA*, 20-26.
England, N. (2009) Mapping values: the vital nature of our uplands. *An atlas linking environment and people*. Policy paper. *Natural England* (NE209).
England, N. (2010) Entry level stewardship: environmental stewardship handbook. Policy paper. *Natural England* (NE349).
Engler, R., Hordijk, W. & Guisan, A. (2012) The MIGCLIM R package–seamless integration of dispersal constraints into projections of species distribution models. *Ecography*, **35,** 872-878.
Erickson, W.P., McDonald, T.L., Gerow, K.G., Howlin, S. & Kern, J.W. (2001) Statistical issues in resource selection studies with radio-marked animals. *Radio Tracking and Animal Populations*, pp. 209-242. Elsevier.
Evans, C.D., Bonn, A., Holden, J., Reed, M.S., Evans, M.G., Worrall, F., Couwenberg, J. & Parnell, M. (2014) Relationships between anthropogenic pressures and ecosystem functions in UK blanket bogs: linking process understanding to ecosystem service valuation. *Ecosystem Services*, **9,** 5-19.
Evans, R. (1997) Soil erosion in the UK initiated by grazing animals: a need for a national survey. *Applied Geography*, **17,** 127-141.
Evans, R. (2005) Reducing soil erosion and the loss of soil fertility for environmentally-sustainable agricultural cropping and livestock production systems. *Annals of Applied Biology*, **146,** 137-146.
Feilhauer, H., He, K.S. & Rocchini, D. (2012) Modeling species distribution using niche-based proxies derived from composite bioclimatic variables and MODIS NDVI. *Remote Sensing*, **4,** 2057-2075.
Fernandez, M., Yesson, C., Gannier, A., Miller, P.I. & Azevedo, J.M. (2017) The importance of temporal resolution for niche modelling in dynamic marine environments. *Journal of Biogeography*, **44,** 2816-2827.
Fieberg, J. (2007) Kernel density estimators of home range: smoothing and the autocorrelation red herring. *Ecology*, **88,** 1059-1066.
Fisher, B., Turner, R.K. & Morling, P. (2009) Defining and classifying ecosystem services for decision making. *Ecological Economics*, **68,** 643-653.
Fisher, G. & Walker, M. (2015) Habitat restoration for curlew *Numenius arquata* at the Lake Vyrnwy reserve, Wales. *Conservation Evidence*, **12,** 48-52.
Fleming, C.H. & Calabrese, J.M. (2017) A new kernel density estimator for accurate home-range and species-range area estimation. *Methods in Ecology and Evolution*, **8,** 571-579.
Fleming, C.H., Calabrese, J.M., Mueller, T., Olson, K.A., Leimgruber, P. & Fagan, W.F. (2014) From fine-scale foraging to home ranges: a semivariance approach to identifying movement modes across spatiotemporal scales. *The American Naturalist*, **183,** E154-E167.
Fleming, C.H., Fagan, W.F., Mueller, T., Olson, K.A., Leimgruber, P. & Calabrese, J.M. (2015) Rigorous home range estimation with movement data: a new autocorrelated kernel density estimator. *Ecology*, **96,** 1182-1188.
Follett, R.F. & Reed, D.A. (2010) Soil carbon sequestration in grazing lands: societal benefits and policy implications. *Rangeland Ecology & Management*, **63,** 4-15.

Forabosco, F., Chitchyan, Z. & Mantovani, R. (2017) Methane, nitrous oxide emissions and mitigation strategies for livestock in developing countries: A review. *South African Journal of Animal Science,* **47,** 268-280.

Forin-Wiart, M.-A., Hubert, P., Sirguey, P. & Poulle, M.-L. (2015) Performance and accuracy of lightweight and low-cost GPS data loggers according to antenna positions, fix intervals, habitats and animal movements. *PLoS One,* **10,** e0129271.

Forsmoo, J., Anderson, K., Macleod, C.J., Wilkinson, M.E. & Brazier, R. (2018) Drone-based structure-from-motion photogrammetry captures grassland sward height variability. *Journal of Applied Ecology,* **55,** 2587-2599.

Franklin, J. (2010a) *Mapping species distributions: spatial inference and prediction*. Cambridge University Press.

Franklin, J. (2010b) Moving beyond static species distribution models in support of conservation biogeography. *Diversity and Distributions,* **16,** 321-330.

Fraser, M., Theobald, V., Griffiths, J., Morris, S. & Moorby, J. (2009) Comparative diet selection by cattle and sheep grazing two contrasting heathland communities. *Agriculture, Ecosystems & Environment,* **129,** 182-192.

Fraser, M.D., Moorby, J.M., Vale, J.E. & Evans, D.M. (2014) Mixed grazing systems benefit both upland biodiversity and livestock production. *PloS one,* **9,** e89054.

Frick, A., Steffenhagen, P., Zerbe, S., Timmermann, T. & Schulz, K. (2011) Monitoring of the vegetation composition in rewetted peatland with iterative decision tree classification of satellite imagery. *Photogrammetrie-Fernerkundung-Geoinformation,* **2011,** 109-122.

Friedman, J., Hastie, T. & Tibshirani, R. (2001) *The elements of statistical learning*. Springer series in statistics New York.

Gahegan, M. (2003) Is inductive machine learning just another wild goose (or might it lay the golden egg)? *International Journal of Geographical Information Science,* **17,** 69-92.

García, R.R., Fraser, M.D., Celaya, R., Ferreira, L.M.M., García, U. & Osoro, K. (2013) Grazing land management and biodiversity in the Atlantic European heathlands: a review. *Agroforestry Systems,* **87,** 19-43.

Gignac, L., Halsey, L. & Vitt, D. (2000) A bioclimatic model for the distribution of Sphagnum-dominated peatlands in North America under present climatic conditions. *Journal of Biogeography,* **27,** 1139-1151.

Girard, I., Ouellet, J.-P., Courtois, R., Dussault, C. & Breton, L. (2002) Effects of sampling effort based on GPS telemetry on home-range size estimations. *The Journal of Wildlife Management,* 1290-1300.

Gobeyn, S., Mouton, A.M., Cord, A.F., Kaim, A., Volk, M. & Goethals, P.L. (2019) Evolutionary algorithms for species distribution modelling: A review in the context of machine learning. *Ecological Modelling,* **392,** 179-195.

Gorham, E. & Rochefort, L. (2003) Peatland restoration: a brief assessment with special reference to Sphagnum bogs. *Wetlands Ecology and Management,* **11,** 109-119.

Grand-Clement, E., Anderson, K., Smith, D., Luscombe, D., Gatis, N., Ross, M. & Brazier, R.E. (2013) Evaluating ecosystem goods and services after restoration of marginal upland peatlands in S outh-W est E ngland. *Journal of Applied Ecology,* **50,** 324-334.

Green, S., Boardman, C., Baird, A. & Gauci, V. (2011) Investigation of peatland restoration (grip blocking) techniques to achieve best outcomes for methane and greenhouse gas emissions/balance. Controlled Environment (Mesocosm) Experiment. Final Report to Defra. Project code SP1202.

Green, S.M., Baird, A.J., Holden, J., Reed, D., Birch, K. & Jones, P. (2017) An experimental study on the response of blanket bog vegetation and water tables to ditch blocking. *Wetlands Ecology and Management,* **25,** 703-716.

Grenouillet, G., Buisson, L., Casajus, N. & Lek, S. (2011) Ensemble modelling of species distribution: the effects of geographical and environmental ranges. *Ecography,* **34,** 9-17.

Grinnell, J. (1917) The niche-relationships of the California Thrasher. *The Auk,* **34,** 427-433.

Guillera-Arroita, G., Lahoz-Monfort, J.J., Elith, J., Gordon, A., Kujala, H., Lentini, P.E., McCarthy, M.A., Tingley, R. & Wintle, B.A. (2015) Is my species distribution model fit for purpose? Matching data and models to applications. *Global Ecology and Biogeography,* **24,** 276-292.
Guisan, A., Edwards Jr, T.C. & Hastie, T. (2002) Generalised linear and generalised additive models in studies of species distributions: setting the scene. *Ecological Modelling,* **157,** 89-100.
Guisan, A. & Theurillat, J.-P. (2000) Equilibrium modeling of alpine plant distribution: how far can we go? *Phytocoenologia,* **30,** 353-384.
Guisan, A. & Thuiller, W. (2005) Predicting species distribution: offering more than simple habitat models. *Ecology Letters,* **8,** 993-1009.
Guisan, A., Thuiller, W. & Zimmermann, N.E. (2017) *Habitat suitability and distribution models: with applications in R*. Cambridge University Press.
Guisan, A. & Zimmermann, N.E. (2000) Predictive habitat distribution models in ecology. *Ecological Modelling,* **135,** 147-186.
Hanley, N., Whitby, M. & Simpson, I. (1999) Assessing the success of agri-environmental policy in the UK. *Land use policy,* **16,** 67-80.
Hanley, T.A. (1982) The nutritional basis for food selection by ungulates. *Rangeland Ecology & Management/Journal of Range Management Archives,* **35,** 146-151.
Hao, T., Elith, J., Lahoz-Monfort, J.J. & Guillera-Arroita, G. (2020) Testing whether ensemble modelling is advantageous for maximising predictive performance of species distribution models. *Ecography,* **43,** 549-558.
Hao, Y. (2019) Characterization of Peat Bog CO_2 and CH_4 Production Potentials in relation to Peat Physico-chemical Properties and Vegetation Composition (unpublished MSc thesis). The Ohio State University.
Harrison, S. (1997) How natural habitat patchiness affects the distribution of diversity in Californian serpentine chaparral. *Ecology,* **78,** 1898-1906.
Hartnett, D.C., Hickman, K.R. & Walter, L. (1996) Effects of bison grazing, fire, and topography on floristic diversity in tallgrass prairie. *Rangeland Ecology & Management/Journal of Range Management Archives,* **49,** 413-420.
Harvey, R.J., Garbutt, A., Hawkins, S.J. & Skov, M.W. (2019) No detectable broad-scale effect of livestock grazing on soil blue-carbon stock in salt marshes. *Frontiers in Ecology and Evolution,* **7,** 151.
Hejcman, M., Češková, M., & Pavlů, V. (2010). Control of Molinia caerulea by cutting management on sub-alpine grassland. *Flora-Morphology, Distribution, Functional Ecology of Plants*, *205*(**9**), 577-582.
Herrero, M., Havlík, P., Valin, H., Notenbaert, A., Rufino, M.C., Thornton, P.K., Blümmel, M., Weiss, F., Grace, D. & Obersteiner, M. (2013) Biomass use, production, feed efficiencies, and greenhouse gas emissions from global livestock systems. *Proceedings of the National Academy of Sciences,* **110,** 20888-20893.
Hijmans, R.J. & van Etten, J. (2016) raster: Geographic data analysis and modeling. *R package version,* **2**.
Hijmans, R.J., Van Etten, J., Cheng, J., Mattiuzzi, M., Sumner, M., Greenberg, J.A., Lamigueiro, O.P., Bevan, A., Racine, E.B. & Shortridge, A. (2015) Package 'raster'. *R package*.
Hinkley, J. (2018) Factors affecting sheep grazing (unpublished MSc thesis). Aberystwyth University.
Hirzel, A. & Guisan, A. (2002) Which is the optimal sampling strategy for habitat suitability modelling. *Ecological Modelling,* **157,** 331-341.
Hirzel, A.H. & Le Lay, G. (2008) Habitat suitability modelling and niche theory. *Journal of Applied Ecology,* **45,** 1372-1381.
Hobbs, R.J., Hallett, L.M., Ehrlich, P.R. & Mooney, H.A. (2011) Intervention ecology: applying ecological science in the twenty-first century. *BioScience,* **61,** 442-450.
Hodgson, J.C., Mott, R., Baylis, S.M., Pham, T.T., Wotherspoon, S., Kilpatrick, A.D., Raja Segaran, R., Reid, I., Terauds, A. & Koh, L.P. (2018) Drones count wildlife more accurately and precisely than humans. *Methods in Ecology and Evolution,* **9,** 1160-1167.

Holden, J., Gascoign, M. & Bosanko, N.R. (2007) Erosion and natural revegetation associated with surface land drains in upland peatlands. *Earth Surface Processes and Landforms,* **32,** 1547-1557.

Hutchinson, G.E. (1957) Cold spring harbor symposium on quantitative biology. *Concluding Remarks,* **22,** 415-427.

IUCN. (2018) A secure peatland future: A vision and strategy for the protection, restoration and sustainable management of UK peatlands. *IUCN UK Peatland Programme, Edinburgh*.

Johnson, C.J. & Gillingham, M.P. (2008) Sensitivity of species-distribution models to error, bias, and model design: an application to resource selection functions for woodland caribou. *Ecological Modelling,* **213,** 143-155.

Johnson, D.S., Thomas, D.L., Ver Hoef, J.M. & Christ, A. (2008) A general framework for the analysis of animal resource selection from telemetry data. *Biometrics,* **64,** 968-976.

Johnson, W.M. (1953) *Effects of Grazing Intensity Upon Vegetation: And Cattle Gains on Ponderosa Pine-bunchgrass Ranges of the Front Range of Colorado*. US Department of Agriculture.

Joosten, H., Gaudig, G., Krawczynski, R., Tanneberger, F., Wichmann, S. & Wichtmann, W. (2014) 25 Managing Soil Carbon in Europe: Paludicultures as a New Perspective for Peatlands. *Soil carbon: Science, management and policy for multiple benefits,* **71,** 297.

Joosten, H., Sirin, A., Couwenberg, J., Laine, J. & Smith, P. (2016) The role of peatlands in climate regulation. *Peatland restoration and ecosystem services: Science, policy and practice,* **66**.

Kang, S., Post, W.M., Nichols, J.A., Wang, D., West, T.O., Bandaru, V. & Izaurralde, R.C. (2013) Marginal lands: concept, assessment and management. *Journal of Agricultural Science,* **5,** 129.

Kays, R., Sheppard, J., Mclean, K., Welch, C., Paunescu, C., Wang, V., Kravit, G. & Crofoot, M. (2019) Hot monkey, cold reality: surveying rainforest canopy mammals using drone-mounted thermal infrared sensors. *International journal of remote sensing,* **40,** 407-419.

Kearney, M. & Porter, W. (2009) Mechanistic niche modelling: combining physiological and spatial data to predict species' ranges. *Ecology letters,* **12,** 334-350.

Kerr, J.T. & Ostrovsky, M. (2003) From space to species: ecological applications for remote sensing. *Trends in ecology & evolution,* **18,** 299-305.

Khatibi, M. & Sheikholeslami, R. (2016) Ecological Niche Theory: A Brief Review. *The International Journal of Indian Psychology,* **3,** 2349-3429.

Kie, J.G., Matthiopoulos, J., Fieberg, J., Powell, R.A., Cagnacci, F., Mitchell, M.S., Gaillard, J.-M. & Moorcroft, P.R. (2010) The home-range concept: are traditional estimators still relevant with modern telemetry technology? *Philosophical Transactions of the Royal Society B: Biological Sciences,* **365,** 2221-2231.

Kimmel, K. & Mander, Ü. (2010) Ecosystem services of peatlands: Implications for restoration. *Progress in Physical Geography,* **34,** 491-514.

Knoth, C., Klein, B., Prinz, T. & Kleinebecker, T. (2013) Unmanned aerial vehicles as innovative remote sensing platforms for high-resolution infrared imagery to support restoration monitoring in cut-over bogs. *Applied Vegetation Science,* **16,** 509-517.

Komulainen, V.M., Tuittila, E.S., Vasander, H. & Laine, J. (1999) Restoration of drained peatlands in southern Finland: initial effects on vegetation change and CO_2 balance. *Journal of Applied Ecology,* **36,** 634-648.

Körner, F., Speck, R., Göktogan, A.H. & Sukkarieh, S. (2010) Autonomous airborne wildlife tracking using radio signal strength. *Intelligent Robots and Systems (IROS), 2010* IEEE/*RSJ International Conference on*, pp. 107-112. IEEE.

Lake, P.S. (2001) On the maturing of restoration: linking ecological research and restoration. *Ecological Management & Restoration,* **2,** 110-115.

Laver, P.N. & Kelly, M.J. (2008) A critical review of home range studies. *The Journal of Wildlife Management,* **72,** 290-298.

Leathwick, J., Elith, J. & Hastie, T. (2006) Comparative performance of generalised additive models and multivariate adaptive regression splines for statistical modelling of species distributions. *Ecological modelling,* **199,** 188-196.

Leathwick, J., Rowe, D., Richardson, J., Elith, J. & Hastie, T. (2005) Using multivariate adaptive regression splines to predict the distributions of New Zealand's freshwater diadromous fish. *Freshwater Biology*, **50,** 2034-2052.

Leibold, M.A. (1995) The niche concept revisited: mechanistic models and community context. *Ecology*, **76,** 1371-1382.

Lennon, M. & Scott, M. (2014) Delivering ecosystems services via spatial planning: reviewing the possibilities and implications of a green infrastructure approach. *Town Planning Review*, **85,** 563-587.

Littlewood, N., Greenwood, S., Quin, S., Pakeman, R. & Woodin, S. (2014) Long-term trends in restored moorland vegetation assemblages. *Community Ecology*, **15,** 104-112.

Lobo, J.M., Jiménez-Valverde, A. & Hortal, J. (2010) The uncertain nature of absences and their importance in species distribution modelling. *Ecography*, **33,** 103-114.

Lovitt, J., Rahman, M.M. & McDermid, G.J. (2017) Assessing the value of UAV photogrammetry for characterizing terrain in complex peatlands. *Remote Sensing*, **9,** 715.

Lucas, R., Medcalf, K., Brown, A., Bunting, P., Breyer, J., Clewley, D., Keyworth, S. & Blackmore, P. (2011) Updating the Phase 1 habitat map of Wales, UK, using satellite sensor data. *ISPRS Journal of Photogrammetry and Remote Sensing*, **66,** 81-102.

Lunt, P., Allott, T., Anderson, P., Buckler, M., Coupar, A., Jones, P., Labadz, J., Worrall, P. & Evans, M. (2010) Peatland restoration. *Scientific Review commissioned by IUCN UK Peatland Programme Commission of Inquiry into Peatland Restoration, Edinburgh, UK.[online] URL: http://www. iucn-ukpeatlandprogramme. org/sites/www. iucn-uk-peatlandprogramme. org/files/Review% 20Peatland% 20Restoration,% 20June*, **202011**.

Luoto, M., Rekolainen, S., Aakkula, J. & Pykälä, J. (2003) Loss of plant species richness and habitat connectivity in grasslands associated with agricultural change in Finland. *AMBIO: A Journal of the Human Environment*, **32,** 447-452.

MA, M.E.A. (2005) Ecosystems and Human Well-Being: Wetlands and Water Synthesis.

Mancini, F., Dubbini, M., Gattelli, M., Stecchi, F., Fabbri, S. & Gabbianelli, G. (2013) Using unmanned aerial vehicles (UAV) for high-resolution reconstruction of topography: The structure from motion approach on coastal environments. *Remote Sensing*, **5,** 6880-6898.

Manly, B., McDonald, L., Thomas, D.L., McDonald, T.L. & Erickson, W.P. (2007) *Resource selection by animals: statistical design and analysis for field studies*. Springer Science & Business Media.

Manzoor, S.A., Griffiths, G. & Lukac, M. (2018) Species distribution model transferability and model grain size–finer may not always be better. *Scientific Reports*, **8,** 1-9.

Marmion, M., Parviainen, M., Luoto, M., Heikkinen, R.K. & Thuiller, W. (2009) Evaluation of consensus methods in predictive species distribution modelling. *Diversity and Distributions*, **15,** 59-69.

Marteinsdóttir, B., Barrio, I.C. & Jónsdóttir, I.S. (2017) Assessing the ecological impacts of extensive sheep grazing in Iceland. *Icelandic Agricultural Sciences*, **30,** 55-72.

Martin, D., Fraser, M.D., Pakeman, R.J. & Moffat, A.M (2013) Natural England Review of Upland Evidence 2012 - Impact of moorland grazing and stocking rates. *Natural England Evidence Review*, **6.**

Maurya, V., Naqvi, S., Joshi, A. & Mittal, J. (2007) Effect of high temperature stress on physiological responses of Malpurs sheep. *Indian Journal of Animal Sciences (India)*.

McCoshum, S.M., Andreoli, S.L., Stenoien, C.M., Oberhauser, K.S. & Baum, K.A. (2016) Species distribution models for natural enemies of monarch butterfly (Danaus plexippus) larvae and pupae: distribution patterns and implications for conservation. *Journal of Insect Conservation*, **20,** 223-237.

McEvoy, J.F., Hall, G.P. & McDonald, P.G. (2016) Evaluation of unmanned aerial vehicle shape, flight path and camera type for waterfowl surveys: disturbance effects and species recognition. *PeerJ*, **4,** e1831.

McFeeters, S.K. (1996) The use of the Normalised Difference Water Index (NDWI) in the delineation of open water features. *International Journal of Remote Sensing*, **17,** 1425-1432.

Menzi, H., Oenema, O., Burton, C., Shipin, O., Gerber, P., Robinson, T. & Franceschini, G. (2010) Impacts of intensive livestock production and manure management on the environment. *Livestock in a Changing Landscape,* **1,** 139-163.

Merow, C., Smith, M.J., Edwards Jr, T.C., Guisan, A., McMahon, S.M., Normand, S., Thuiller, W., Wüest, R.O., Zimmermann, N.E. & Elith, J. (2014) What do we gain from simplicity versus complexity in species distribution models? *Ecography,* **37,** 1267-1281.

Milligan, A., Putwain, P., Cox, E., Ghorbani, J., Le Duc, M. & Marrs, R. (2004) Developing an integrated land management strategy for the restoration of moorland vegetation on Molinia caerulea-dominated vegetation for conservation purposes in upland Britain. *Biological Conservation,* **119,** 371-385.

Montgomery, D.C., Peck, E.A. & Vining, G.G. (2012) *Introduction to linear regression analysis*. John Wiley & Sons.

Moore, P. (1987) Ecological and hydrological aspects of peat formation. *Geological Society, London, Special Publications,* **32,** 7-15.

Moudrý, V. & Šímová, P. (2012) Influence of positional accuracy, sample size and scale on modelling species distributions: a review. *International Journal of Geographical Information Science,* **26,** 2083-2095.

Naimi, B. & Araújo, M.B. (2016) sdm: a reproducible and extensible R platform for species distribution modelling. *Ecography,* **39,** 368-375.

Naimi, B., Hamm, N.A., Groen, T.A., Skidmore, A.K. & Toxopeus, A.G. (2014) Where is positional uncertainty a problem for species distribution modelling? *Ecography,* **37,** 191-203.

National Trust. (2015). Managing Molinia? *Proceedings of a 3-day conference 14-16 September 2015,* Huddersfield, West Yorkshire, UK.

Nilsen, E.B., Pedersen, S. & Linnell, J.D. (2008) Can minimum convex polygon home ranges be used to draw biologically meaningful conclusions? *Ecological research,* **23,** 635-639.

Noble, A. (2018) An evaluation of sheep movement and behaviour in Wales (Unpublished MSc thesis). Aberystwyth University.

Odland, J. (1988) *Spatial autocorrelation*. SAGE Publications, Incorporated.

Oke, T.A. & Hager, H.A. (2017) Assessing environmental attributes and effects of climate change on Sphagnum peatland distributions in North America using single-and multi-species models. *PloS one,* **12,** e0175978.

Oksanen, J. (1997) Why the beta-function cannot be used to estimate skewness of species responses. *Journal of Vegetation Science,* **8,** 147-152.

Oksanen, J. & Minchin, P.R. (2002) Continuum theory revisited: what shape are species responses along ecological gradients? *Ecological Modelling,* **157,** 119-129.

Orians, G.H. & Wittenberger, J.F. (1991) Spatial and temporal scales in habitat selection. *The American Naturalist,* **137,** S29-S49.

Orr, H.G., Wilby, R., Hedger, M.M. & Brown, I. (2008) Climate change in the uplands: a UK perspective on safeguarding regulatory ecosystem services. *Climate Research,* **37,** 77-98.

Owen-Smith, N. & Novellie, P. (1982) What should a clever ungulate eat? *The American Naturalist,* **119,** 151-178.

Palace, M., Herrick, C., DelGreco, J., Finnell, D., Garnello, A.J., McCalley, C., McArthur, K., Sullivan, F. & Varner, R.K. (2018) Determining subarctic peatland vegetation using an unmanned aerial system (UAS). *Remote Sensing,* **10,** 1498.

Parry, L.E., Holden, J. & Chapman, P.J. (2014) Restoration of blanket peatlands. *Journal of Environmental Management,* **133,** 193-205.

Peacock, M., Jones, T., Airey, B., Johncock, A., Evans, C., Lebron, I., Fenner, N. & Freeman, C. (2015) The effect of peatland drainage and rewetting (ditch blocking) on extracellular enzyme activities and water chemistry. *Soil Use and Management,* **31,** 67-76.

Pearson, R.G. (2007) Species' distribution modeling for conservation educators and practitioners. *Synthesis. American Museum of Natural History,* **50,** 54-89.

Pebesma, E., Bivand, R., Pebesma, M.E., RColorBrewer, S. & Collate, A. (2012) Package 'sp'. *The Comprehensive R Archive Network.*

Peek, J.M., Scott, M.D., Nelson, L.J., Pierce, D.J. & Irwin, L.L. (1982) Role of cover in habitat management for big game in northwestern United-States. *Transactions of the North American Wildlife and Natural Resources Conference*, pp. 363-373. White Horse Press, 1 Strond, Isle of Harris, Scotland.

Pegard, A., Miquel, C., Valentini, A., Coissac, E., Bouvier, F., Francois, D., Taberlet, P., Engel, E. & Pompanon, F. (2009) Universal DNA-based methods for assessing the diet of grazing livestock and wildlife from feces. *Journal of Agricultural and Food Chemistry,* **57,** 5700-5706.

Perotto-Baldivieso, H.L., Cooper, S.M., Cibils, A.F., Figueroa-Pagán, M., Udaeta, K. & Black-Rubio, C.M. (2012) Detecting autocorrelation problems from GPS collar data in livestock studies. *Applied Animal Behaviour Science,* **136,** 117-125.

Peterson, A.T., Papeş, M. & Soberón, J. (2015) Mechanistic and correlative models of ecological niches. *European Journal of Ecology,* **1,** 28-38.

Peterson, A.T. & Soberón, J. (2012) Species distribution modeling and ecological niche modeling: getting the concepts right. *Natureza & Conservação,* **10,** 102-107.

Pettorelli, N., Vik, J.O., Mysterud, A., Gaillard, J.-M., Tucker, C.J. & Stenseth, N.C. (2005) Using the satellite-derived NDVI to assess ecological responses to environmental change. *Trends in Ecology & Evolution,* **20,** 503-510.

Pix4D. (2020). Radiometric corrections. Retrieved from https://support.pix4d.com/hc/en-us/articles/202559509-Radiometric-corrections

Porfirio, L.L., Harris, R.M., Lefroy, E.C., Hugh, S., Gould, S.F., Lee, G., Bindoff, N.L. & Mackey, B. (2014) Improving the use of species distribution models in conservation planning and management under climate change. *PLoS One,* **9,** e113749.

Pradhan, P. (2016) Strengthening MaxEnt modelling through screening of redundant explanatory bioclimatic variables with variance inflation factor analysis. *Researcher,* **8,** 29-34.

Pulliam, H.R. (2000) On the relationship between niche and distribution. *Ecology letters,* **3,** 349-361.

Putfarken, D., Dengler, J., Lehmann, S. & Härdtle, W. (2008) Site use of grazing cattle and sheep in a large-scale pasture landscape: a GPS/GIS assessment. *Applied Animal Behaviour Science,* **111,** 54-67.

Péron, G. (2019) The time frame of home-range studies: from function to utilization. *Biological Reviews,* **94,** 1974-1982.

Rahman, M.M., McDermid, G.J., Strack, M. & Lovitt, J. (2017) A new method to map groundwater table in peatlands using unmanned aerial vehicles. *Remote Sensing,* **9,** 1057.

Raynolds, M., Magnússon, B., Metúsalemsson, S. & Magnússon, S.H. (2015) Warming, sheep and volcanoes: Land cover changes in Iceland evident in satellite NDVI trends. *Remote Sensing,* **7,** 9492-9506.

Reed, M., Arblaster, K., Bullock, C., Burton, R., Davies, A., Holden, J., Hubacek, K., May, R., Mitchley, J. & Morris, J. (2009) Using scenarios to explore UK upland futures. *Futures,* **41,** 619-630.

Renner, I.W. & Warton, D.I. (2013) Equivalence of MAXENT and Poisson point process models for species distribution modeling in ecology. *Biometrics,* **69,** 274-281.

Robertson, M.P., Peter, C.I., Villet, M.H. & Ripley, B.S. (2003) Comparing models for predicting species' potential distributions: a case study using correlative and mechanistic predictive modelling techniques. *Ecological Modelling,* **164,** 153-167.

Robertson, M.P., Villet, M.H. & Palmer, A.R. (2004) A fuzzy classification technique for predicting species' distributions: applications using invasive alien plants and indigenous insects. *Diversity and Distributions,* **10,** 461-474.

Rosenburgh, A.E. (2015) Restoration and recovery of Sphagnum on degraded blanket bog. Manchester Metropolitan University.

RSPB (2007) *The Uplands.* Time to Change? Sandy, Bedfordshire.

Ryan, J.C., Hubbard, A.L., Box, J.E., Todd, J., Christoffersen, P., Carr, J.R., Holt, T.O. & Snooke, N.A. (2015) UAV photogrammetry and structure from motion to assess calving dynamics at Store Glacier, a large outlet draining the Greenland ice sheet. *The Cryosphere*, **9**, 1–11,

Rykiel Jr, E.J. (1996) Testing ecological models: the meaning of validation. *Ecological modelling,* **90,** 229-244.
Rümmler, M.-C., Mustafa, O., Maercker, J., Peter, H.-U. & Esefeld, J. (2018) Sensitivity of Adélie and Gentoo penguins to various flight activities of a micro UAV. *Polar Biology,* **41,** 2481-2493.
Segurado, P. & Araujo, M.B. (2004) An evaluation of methods for modelling species distributions. *Journal of biogeography,* **31,** 1555-1568.
Shafer, M.W., Vega, G., Rothfus, K. & Flikkema, P. (2019) UAV wildlife radio telemetry: System and methods of localization. *Methods in Ecology and Evolution*.
Shannon, N.H., Hudson, R.J., Brink, V.C. & Kitts, W.D. (1975) Determinants of spatial distribution of Rocky Mountain bighorn sheep. *The Journal of Wildlife Management***,** 387-401.
Silcock, P., Brunyee, J. & Pring, J. (2012) Changing livestock numbers in the UK Less Favoured Areas-an anlaysis of likely biodiversity implications. Final Report prepared for the Royal Society for the Protection of Birds.
Sillero, N. (2011) What does ecological modelling model? A proposed classification of ecological niche models based on their underlying methods. *Ecological Modelling,* **222,** 1343-1346.
Stahl, P., Vandel, J., Ruette, S., Coat, L., Coat, Y. & Balestra, L. (2002) Factors affecting lynx predation on sheep in the French Jura. *Journal of Applied Ecology,* **39,** 204-216.
Stevens, E., Stevens, S.J., Gates, R., Eskridge, K. & Waller, S. (1987) Procedure for fecal cuticle analysis of herbivore diets. *Rangeland Ecology & Management/Journal of Range Management Archives,* **40,** 187-189.
Stevens, J., Blackstock, T., Howe, E. & Stevens, D. (2004) Repeatability of Phase 1 habitat survey. *Journal of Environmental Management,* **73,** 53-59.
Stockwell, D.R. & Peterson, A.T. (2002) Effects of sample size on accuracy of species distribution models. *Ecological modelling,* **148,** 1-13.
Stolar, J. & Nielsen, S.E. (2015) Accounting for spatially biased sampling effort in presence-only species distribution modelling. *Diversity and Distributions,* **21,** 595-608.
Svenning, J.C., Normand, S. & Kageyama, M. (2008) Glacial refugia of temperate trees in Europe: insights from species distribution modelling. *Journal of Ecology,* **96,** 1117-1127.
Swain, D.L., Wark, T. & Bishop-Hurley, G. (2008) Using high fix rate GPS data to determine the relationships between fix rate, prediction errors and patch selection. *Ecological Modelling,* **212,** 273-279.
Szabó, G., Singh, S.K. & Szabó, S. (2015) Slope angle and aspect as influencing factors on the accuracy of the SRTM and the ASTER GDEM databases. *Physics and Chemistry of the Earth, Parts A/B/C,* **83,** 137-145.
Team, Q.D. (2016) QGIS geographic information system. *Open source geospatial Foundation project*.
Team, R.C. (2013) R: A language and environment for statistical computing.
Thuiller, W., Albert, C., Araújo, M.B., Berry, P.M., Cabeza, M., Guisan, A., Hickler, T., Midgley, G.F., Paterson, J. & Schurr, F.M. (2008) Predicting global change impacts on plant species' distributions: future challenges. *Perspectives in Plant Ecology, Evolution and Systematics,* **9,** 137-152.
Thuiller, W., Lafourcade, B., Engler, R. & Araújo, M.B. (2009) BIOMOD–a platform for ensemble forecasting of species distributions. *Ecography,* **32,** 369-373.
Torres, L.G., Orben, R.A., Tolkova, I. & Thompson, D.R. (2017) Classification of animal movement behavior through residence in space and time. *PloS one,* **12,** e0168513.
Tranter, R.B., Swinbank, A., Wooldridge, M., Costa, L., Knapp, T., Little, G.J. & Sottomayor, M.L. (2007) Implications for food production, land use and rural development of the European Union's Single Farm Payment: Indications from a survey of farmers' intentions in Germany, Portugal and the UK. *Food policy,* **32,** 656-671.
Tremblay, J.A., Desrochers, A., Aubry, Y., Pace, P. & Bird, D.M. (2017) A low-cost technique for radio-tracking wildlife using a small standard unmanned aerial vehicle. *Journal of Unmanned Vehicle Systems,* **5,** 102-108.

Tuohy, A., Bazilian, M., Doherty, R., Gallachóir, B.Ó. & O'Malley, M. (2009) Burning peat in Ireland: An electricity market dispatch perspective. *Energy Policy,* **37,** 3035-3042.
Ungar, E.D., Henkin, Z., Gutman, M., Dolev, A., Genizi, A. & Ganskopp, D. (2005) Inference of animal activity from GPS collar data on free-ranging cattle. *Rangeland Ecology & Management,* **58,** 256-266.
Vas, E., Lescroël, A., Duriez, O., Boguszewski, G. & Grémillet, D. (2015) Approaching birds with drones: first experiments and ethical guidelines. *Biology letters,* **11,** 20140754.
Vavra, M. & Ganskopp, D. (1987) Slope use by cattle, feral horses, deer, and bighorn sheep.
Walter, W.D., Onorato, D.P. & Fischer, J.W. (2015) Is there a single best estimator? Selection of home range estimators using area-under-the-curve. *Movement Ecology,* **3,** 10.
Ward, S.E., Bardgett, R.D., McNamara, N.P., Adamson, J.K. & Ostle, N.J. (2007) Long-term consequences of grazing and burning on northern peatland carbon dynamics. *Ecosystems,* **10,** 1069-1083.
Whittaker, R.H. (1956) Vegetation of the great smoky mountains. *Ecological Monographs,* **26,** 1-80.
Wickham, H., Francois, R., Henry, L. & Müller, K. (2015) dplyr: A grammar of data manipulation. *R package version 0.4,* **3**.
Williamson, J., Rowe, E., Reed, D., Ruffino, L., Jones, P., Dolan, R., Buckingham, H., Norris, D., Astbury, S. & Evans, C.D. (2017) Historical peat loss explains limited short-term response of drained blanket bogs to rewetting. *Journal of Environmental Management,* **188,** 278-286.
Wilson, L., Wilson, J., Holden, J., Johnstone, I., Armstrong, A. & Morris, M. (2010) Recovery of water tables in Welsh blanket bog after drain blocking: discharge rates, time scales and the influence of local conditions. *Journal of Hydrology,* **391,** 377-386.
Worrall, F., Chapman, P., Holden, J., Evans, C., Artz, R., Smith, P. & Grayson, R. (2010) Peatlands and climate change. *Report to IUCN UK Peatland Programme, Edinburgh. www. iucn-ukpeatlandprogramme. org/scientificreviews*.
Worton, B.J. (1989) Kernel methods for estimating the utilization distribution in home-range studies. *Ecology,* **70,** 164-168.
Xintu, L. (2008) Conditions of peat formation. *Encyclopedia of Life Support Systems (EOLSS),* **2,** 11.
Yearby, M. (2018) "Where are my sheep?": Sheep detection in aerial photography through template recognition (unpublished undergraduate thesis). Aberystwyth University.
Yousef, M.A. & Ragheb, M.K. (2014) Effect of Recording Interval on GPS Accuracy. *Journal of Engineering Sciences, Assiut University,* **42,** 1215-1231.
Zimmermann, N.E., Edwards Jr, T.C., Graham, C.H., Pearman, P.B. & Svenning, J.C. (2010) New trends in species distribution modelling. *Ecography,* **33,** 985-989.
Zuckerberg, B., Fink, D., La Sorte, F.A., Hochachka, W.M. & Kelling, S. (2016) Novel seasonal land cover associations for eastern North American forest birds identified through dynamic species distribution modelling. *Diversity and Distributions,* **22,** 717-730.
Zurrell, D. (2019) Introduction to species distribution modelling (SDM) in R. damariszurrell.github.io.

APPENDICES

A.1 Ffridd Fawr stitching report

Summary

Project	Ffridd_fawr
Processed	2020-04-08 15:55:14
Camera Model Name(s)	RedEdge-M_5.5_1280x960 (Blue), RedEdge-M_5.5_1280x960 (Green), RedEdge-M_5.5_1280x960 (Red), RedEdge-M_5.5_1280x960 (NIR), RedEdge-M_5.5_1280x960 (Red edge)
Rig name(s)	«RedEdge-M»
Average Ground Sampling Distance (GSD)	7.18 cm /2.83 in
Area Covered	1.557 km^2 /155.6579 ha /0.60 sq. mi. /384.8382 acres
Time for Initial Processing (without report)	1d:00h:45m:43s

Quality Check

Images	median of 45452 keypoints per image	✓
Dataset	11815 out of 12250 images calibrated (96%), all images enabled, 3 blocks	⚠
Camera Optimization	0.03% relative difference between initial and optimised internal camera parameters	✓
Matching	median of 24021.9 matches per calibrated image	✓
Georeferencing	yes, no 3D GCP	⚠

Preview

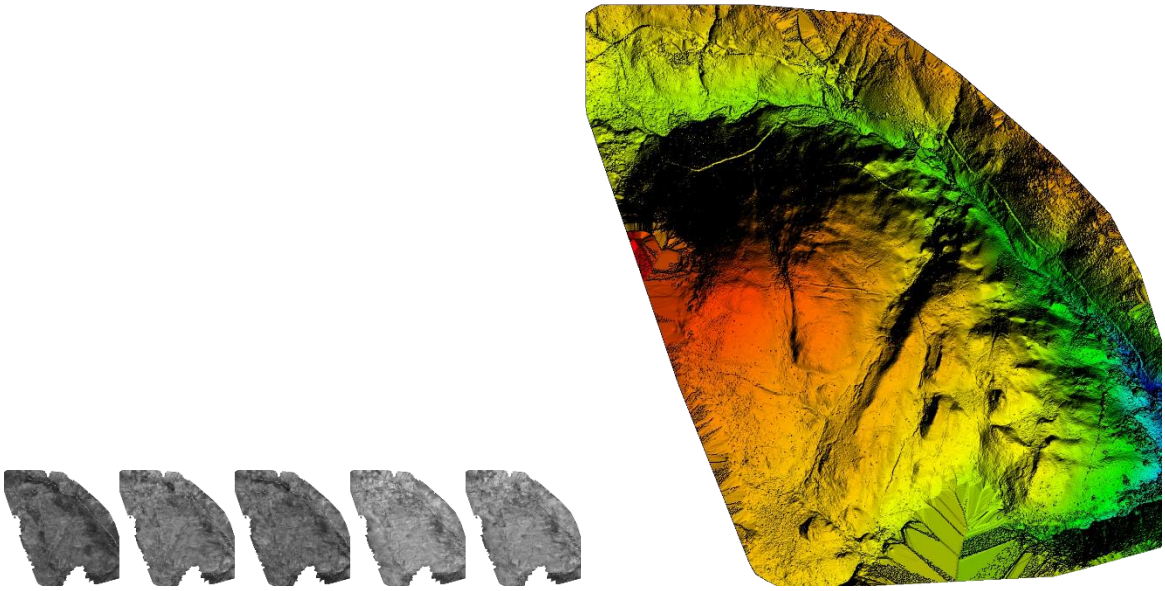


Figure 1: Orthomosaic and the corresponding sparse Digital Surface Model (DSM) before densification.

Calibration Details

Number of Calibrated Images	11815 out of 12250
Number of Geolocated Images	12250 out of 12250

Initial Image Positions

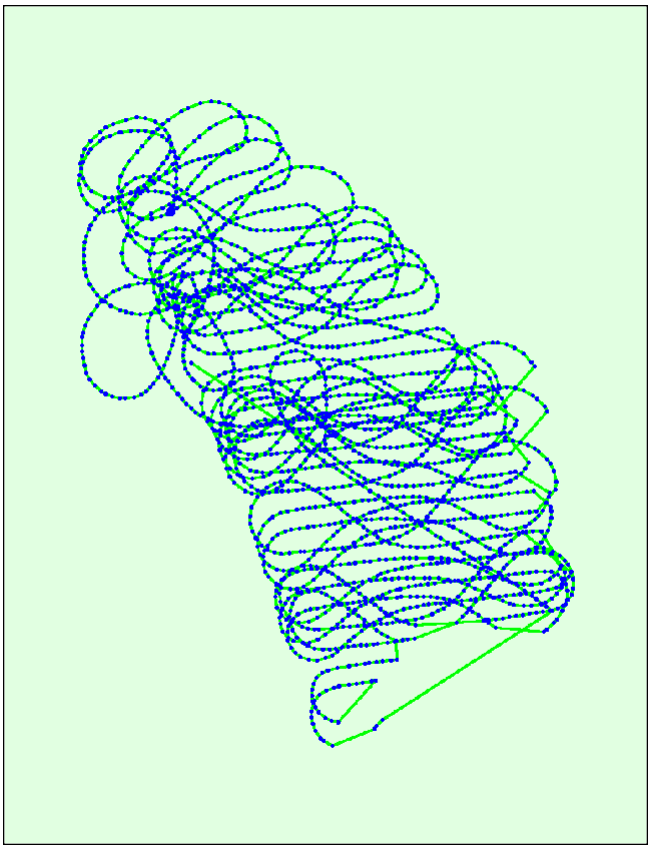
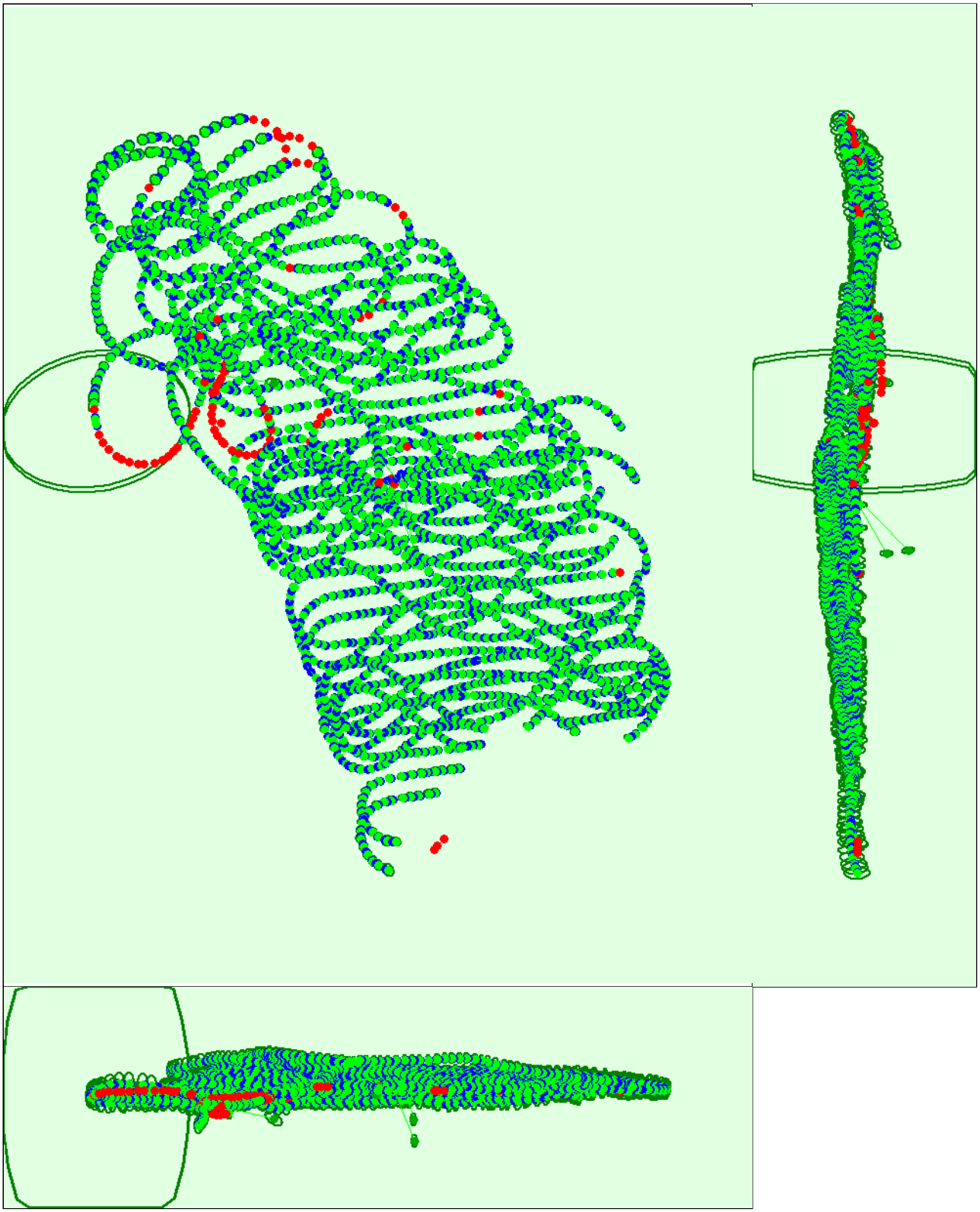


Figure 2: Top view of the initial image position. The green line follows the position of the images in time starting from the large blue dot.

Computed Image/GCPs/Manual Tie Points Positions



Uncertainty ellipses 100 x magnified

Figure 3: Offset between initial (blue dots) and computed (green dots) image positions as well as the offset between the GCPs initial positions (blue crosses) and their computed positions (green crosses) in the top-view (XY plane), front-view (XZ plane), and side-view (YZ plane). Red dots indicate disabled or uncalibrated images. Dark green ellipses indicate the absolute position uncertainty of the bundle block adjustment result.

Absolute camera position and orientation uncertainties

	X[m]	Y[m]	Z [m]	Omega [degree]	Phi [degree]	Kappa [degree]
Mean	0.060	0.059	0.133	116.314	309.766	131.945
Sigma	0.085	0.062	0.172	2364.595	6307.042	2696.590

Overlap

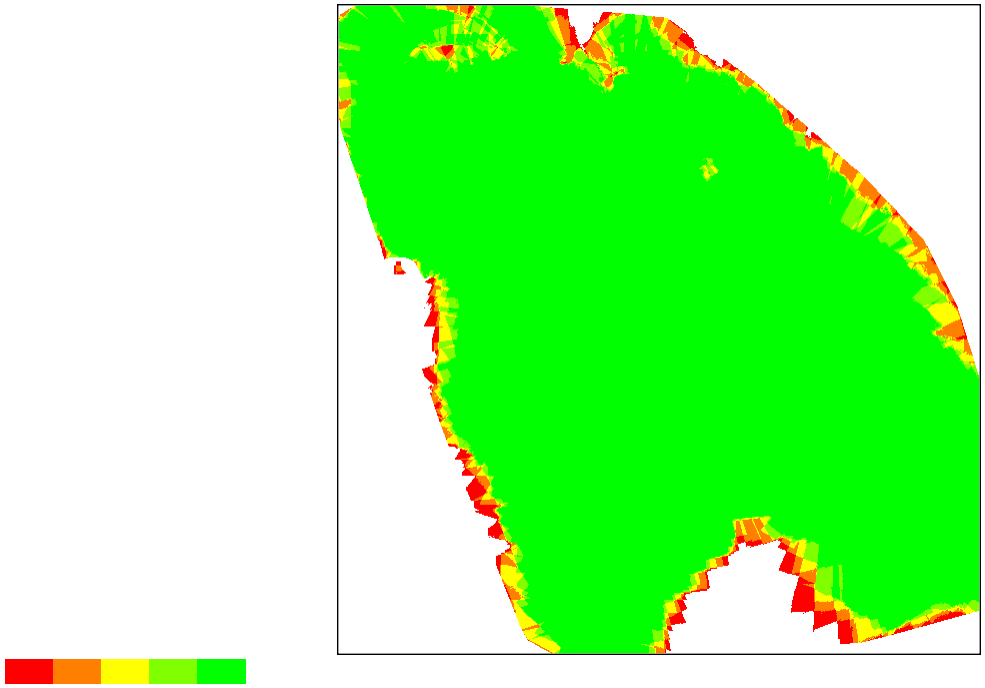


Figure 4: Number of overlapping images computed for each pixel of the orthomosaic.
Red and yellow areas indicate low overlap for which poor results may be generated. Green areas indicate an overlap of over 5 images for every pixel. Good quality results will be generated as long as the number of keypoint matches is also sufficient for these areas (see Figure 5 for keypoint matches).

Bundle Block Adjustment Details

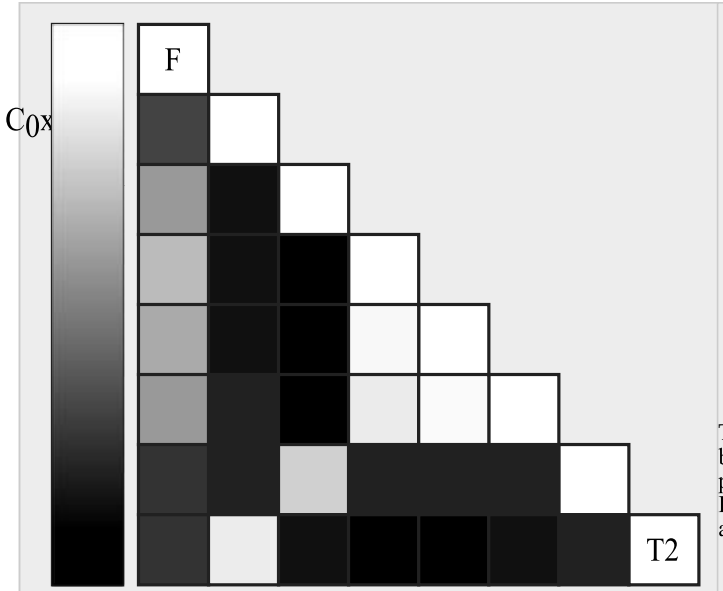
Number of 2D Keypoint Observations for Bundle Block Adjustment	72606873
Number of 3D Points for Bundle Block Adjustment	23308847
Mean Reprojection Error [pixels]	0.147

Internal Camera Parameters

RedEdge-M_5.5_1280x960 (Blue). Sensor Dimensions: 4.800 [mm] x 3.600 [mm]

EXIF ID: RedEdge-M_5.5_1280x960

	Focal Length	Principal Point x	Principal Point y	R1	R2	R3	T1	T2
Initial Values	1447.570 [pixel] 5.428 [mm]	626.181 [pixel] 2.348 [mm]	481.336 [pixel] 1.805 [mm]	-0.099	0.158	-0.063	-0.001	-0.000
Optimised Values	1447.164 [pixel] 5.427 [mm]	624.278 [pixel] 2.341 [mm]	480.920 [pixel] 1.803 [mm]	-0.099	0.154	-0.041	-0.001	-0.000
Uncertainties (Sigma)	0.043 [pixel] 0.000 [mm]	0.085 [pixel] 0.000 [mm]	0.064 [pixel] 0.000 [mm]	0.001	0.004	0.010	0.000	0.000



The correlation between camera internal parameters determined by the bundle adjustment. White indicates a full correlation between the parameters, ie. any change in one can be fully compensated by the other. Black indicates that the parameter is completely independent, and is not affected by other parameters.

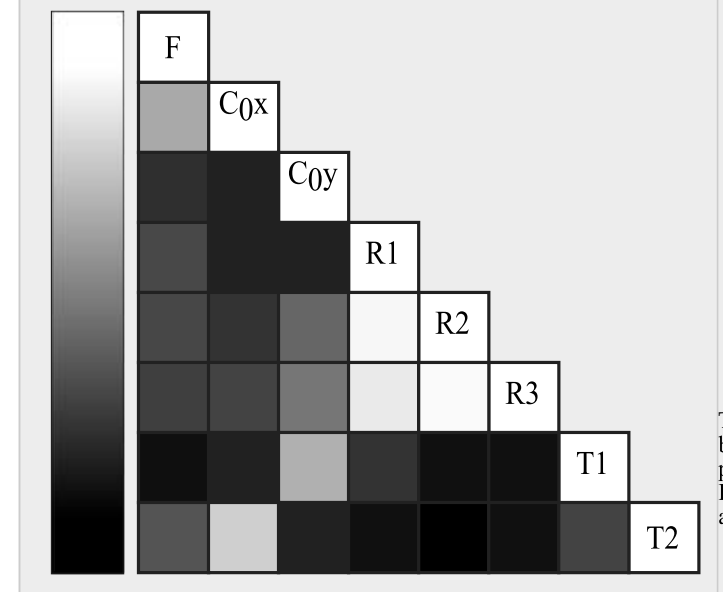
The number of Automatic Tie Points (ATPs) per pixel, averaged over all images of the camera model, is color coded between black and white. White indicates that, on average, more than 16 ATPs have been extracted at the pixel location. Black indicates that, on average, 0 ATPs have been extracted at the pixel location. Click on the image to the see the average direction and magnitude of the re- projection error for each pixel. Note that the vectors are scaled for better visualization. The scale bar indicates the magnitude of 1 pixel error.

Internal Camera Parameters

RedEdge-M_5.5_1280x960 (Green). Sensor Dimensions: 4.800 [mm] x 3.600 [mm]

EXIF ID: RedEdge-M_5.5_1280x960

	Focal Length	Principal Pointx	Principal Pointy	R1	R2	R3	T1	T2
Initial Values	1453.098 [pixel] 5.449 [mm]	621.848 [pixel] 2.332 [mm]	490.765 [pixel] 1.840 [mm]	-0.095	0.134	-0.003	0.000	-0.000
Optimized Values	1452.367 [pixel] 5.446 [mm]	619.536 [pixel] 2.323 [mm]	490.489 [pixel] 1.839 [mm]	-0.096	0.139	-0.014	0.000	-0.001
				0.000	0.001	0.002	0.000	0.000



The correlation between camera internal parameters determined by the bundle adjustment. White indicates a full correlation between the parameters, ie. any change in one can be fully compensated by the other. Black indicates that the parameter is completely independent, and is not affected by other parameters.

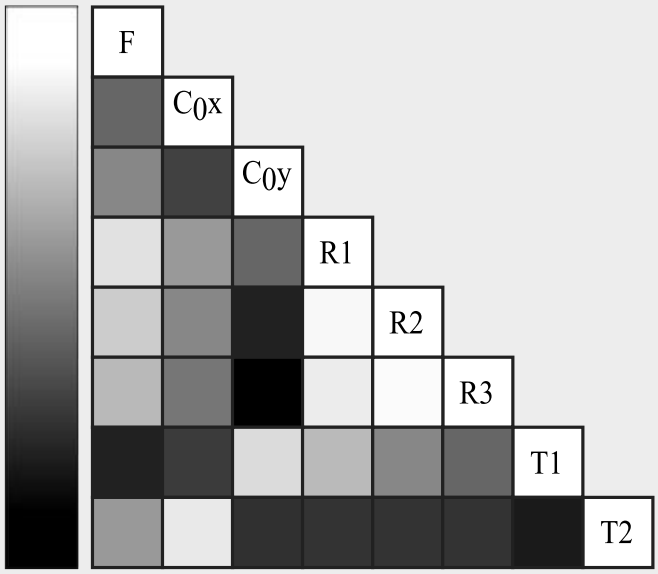
The number of Automatic Tie Points (ATPs) per pixel, averaged over all images of the camera model, is color coded between black and white. White indicates that, on average, more than 16 ATPs have been extracted at the pixel location. Black indicates that, on average, 0 ATPs have been extracted at the pixel location. Click on the image to the see the averaged direction and magnitude of there- projection error for each pixel. Note that the vectors are scaled for better visualization. The scale bar indicates the magnitude of 1 pixel error.

Internal Camera Parameters

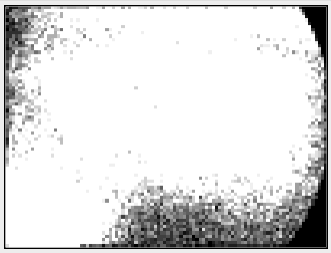
RedEdge-M_5.5_1280x960 (Red). Sensor Dimensions: 4.800 [mm] x 3.600 [mm]

EXIF ID: RedEdge-M_5.5_1280x960

	Focal Length	Principal Point x	Principal Point y	R1	R2	R3	T1	T2
Initial Values	1446.130 [pixel] 5.423 [mm]	616.776 [pixel] 2.313 [mm]	480.909 [pixel] 1.803 [mm]	-0.104	0.154	-0.049	0.001	-0.000
Optimised Values	1445.294 [pixel] 5.420 [mm]	612.975 [pixel] 2.299 [mm]	480.234 [pixel] 1.801 [mm]	-0.103	0.151	-0.042	0.001	-0.001
Uncertainties (Sigma)	0.087 [pixel] 0.000 [mm]	0.194 [pixel] 0.001 [mm]	0.165 [pixel] 0.001 [mm]	0.001	0.010	0.022	0.000	0.000



The correlation between camera internal parameters determined by the bundle adjustment. White indicates a full correlation between the parameters, ie. any change in one can be fully compensated by the other. Black indicates that the parameter is completely independent, and is not affected by other parameters.



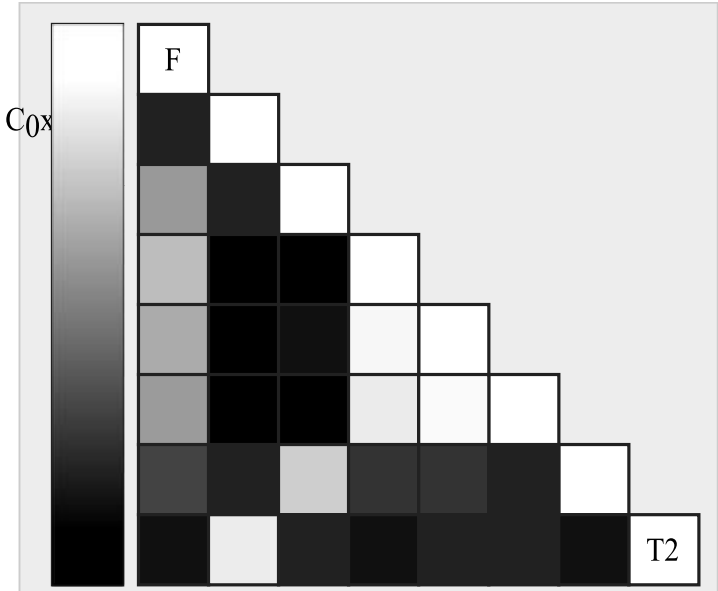
The number of Automatic Tie Points (ATPs) per pixel, averaged over all images of the camera model, is color coded between black and white. White indicates that, on average, more than 16 ATPs have been extracted at the pixel location. Black indicates that, on average, 0 ATPs have been extracted at the pixel location. Click on the image to the see the average direction and magnitude of the re- projection error for each pixel. Note that the vectors are scaled for better visualization. The scale bar indicates the magnitude of 1 pixel error.

Internal Camera Parameters

RedEdge-M_5.5_1280x960 (NIR). Sensor Dimensions: 4.800 [mm] x 3.600 [mm]

EXIF ID: RedEdge-M_5.5_1280x960

	Focal Length	Principal Point x	Principal Point y	R1	R2	R3	T1	T2
Initial Values	1450.269 [pixel] 5.439 [mm]	633.784 [pixel] 2.377 [mm]	474.400 [pixel] 1.779 [mm]	-0.105	0.144	-0.039	-0.000	-0.000
Optimised Values	1449.965 [pixel] 5.437 [mm]	631.541 [pixel] 2.368 [mm]	474.174 [pixel] 1.778 [mm]	-0.109	0.177	-0.104	-0.000	-0.000
Uncertainties (Sigma)	0.044 [pixel] 0.000 [mm]	0.091 [pixel] 0.000 [mm]	0.068 [pixel] 0.000 [mm]	0.001	0.004	0.010	0.000	0.000



The correlation between camera internal parameters determined by the bundle adjustment. White indicates a full correlation between the parameters, ie. any change in one can be fully compensated by the other. Black indicates that the parameter is completely independent, and is not affected by other parameters.

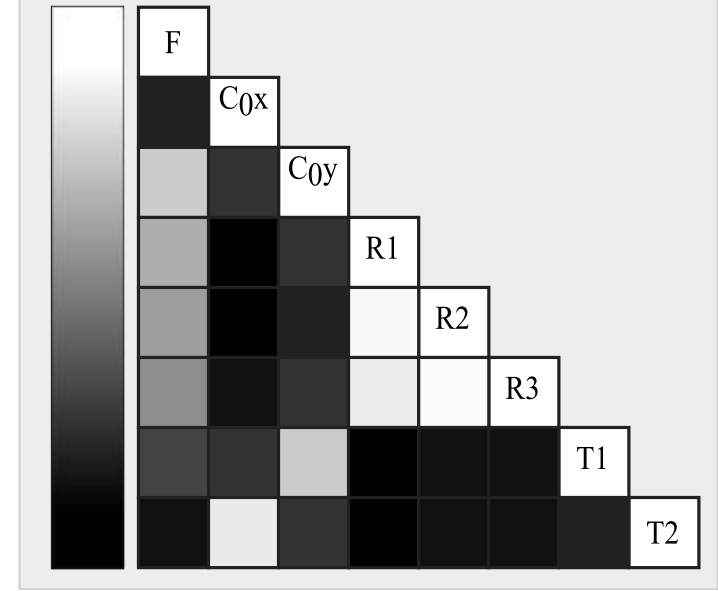
The number of Automatic Tie Points (ATPs) per pixel, averaged over all images of the camera model, is color coded between black and white. White indicates that, on average, more than 16 ATPs have been extracted at the pixel location. Black indicates that, on average, 0 ATPs have been extracted at the pixel location. Click on the image to the see the average direction and magnitude of the re- projection error for each pixel. Note that the vectors are scaled for better visualization. The scale bar indicates the magnitude of 1 pixel error.

Internal Camera Parameters

RedEdge-M_5.5_1280x960 (Red edge). Sensor Dimensions: 4.800 [mm] x 3.600 [mm]

EXIF ID: RedEdge-M_5.5_1280x960

	Focal Length	Principal Point x	Principal Point y	R1	R2	R3	T1	T2
Initial Values	1446.905 [pixel]	635.379 [pixel]	463.811 [pixel]	-0.103	0.139	-0.021	0.001	0.000
	5.426 [mm]	2.383 [mm]	1.739 [mm]					
Optimized Values	1446.594 [pixel]	633.696 [pixel]	463.681 [pixel]	-0.104	0.149	-0.035	0.001	-0.000
	5.425 [mm]	2.376 [mm]	1.739 [mm]	0.000	0.004	0.008	0.000	0.000



The correlation between camera internal parameters determined by the bundle adjustment. White indicates a full correlation between the parameters, ie. any change in one can be fully compensated by the other. Black indicates that the parameter is completely independent, and is not affected by other parameters.

The number of Automatic Tie Points (ATPs) per pixel, averaged over all images of the camera model, is color coded between black and white. White indicates that, on average, more than 16 ATPs have been extracted at the pixel location. Black indicates that, on average, 0 ATPs have been extracted at the pixel location. Click on the image to the see the averaged direction and magnitude of the re- projection error for each pixel. Note that the vectors are scaled for better visualization. The scale bar indicates the magnitude of 1 pixel error.

Camera Rig «RedEdge-M» Relatives. Images: 12250

	Transl X[m]	Transl Y[m]	TranslZ[m]	Rot X[degree]	Rot Y[degree]	RotZ[degree]
RedEdge-M_5.5_1280x960 (Green)	Reference Camera					
RedEdge-M_5.5_1280x960 (Blue)						
Initial Values	0.030	0.000	0.000	0.000	0.000	0.000
Optimised values	0.030	0.000	0.000	-0.040	-0.016	-0.062
Uncertainties (sigma)				0.003	0.003	0.000
RedEdge-M_5.5_1280x960 (Red)						
Initial Values	0.000	0.021	0.000	0.000	0.000	0.000
Optimised values	0.000	0.021	0.000	0.007	0.190	-0.120
Uncertainties (sigma)				0.007	0.008	0.001
RedEdge-M_5.5_1280x960 (NIR)						
Initial Values	0.030	0.021	0.000	0.000	0.000	0.000
Optimised values	0.030	0.021	0.000	-0.131	0.003	-0.030
Uncertainties (sigma)				0.003	0.004	0.000
RedEdge-M_5.5_1280x960 (Red edge)						
Initial Values	0.015	0.011	0.000	0.000	0.000	0.000
Optimised values	0.015	0.011	0.000	0.040	0.103	0.039
Uncertainties (sigma)				0.002	0.003	0.000

2D Keypoints Table

	Number of 2D Keypoints per Image	Number of Matched 2D Keypoints per Image
Median	45452	24022
Min	18550	0
Max	55580	40538
Mean	43379	22035

2D Keypoints Table for Camera RedEdge-M_5.5_1280x960 (Blue)

	Number of 2D Keypoints per Image	Number of Matched 2D Keypoints per Image
Median	42740	15448
Min	18550	11
Max	51603	28613
Mean	40600	14251

2D Keypoints Table for Camera RedEdge-M_5.5_1280x960 (Green)

	Number of 2D Keypoints per Image	Number of Matched 2D Keypoints per Image
Median	46102	26733
Min	19536	60
Max	55580	40538
Mean	43818	24733

2D Keypoints Table for Camera RedEdge-M_5.5_1280x960 (Red)

	Number of 2D Keypoints per Image	Number of Matched 2D Keypoints per Image
Median	44149	7537

Min	19672	0
Max	49216	30558
Mean	42437	9676

2D Keypoints Table for Camera RedEdge-M_5.5_1280x960 (NIR)

	Number of 2D Keypoints per Image	Number of Matched 2D Keypoints per Image
Median	46159	20184
Min	20076	8
Max	53057	33826
Mean	42945	18696

2D Keypoints Table for Camera RedEdge-M_5.5_1280x960 (Red edge)

	Number of 2D Keypoints per Image	Number of Matched 2D Keypoints per Image
Median	46101	19812
Min	19364	13
Max	52393	32491
Mean	43086	18165

Median /75%/Maximal Number of Matches Between Camera Models

	RedEdge-M_5.5_... (Blue)	RedEdge-M_5.5... (Green)	RedEdge-M_5.5_1... (Red)	RedEdge-M_5.5_1...(NIR)	RedEdge-M_...(Red edge)
RedEdge-M_5.5_1280x960 (Blue)	180/2155/19307	96 /518 /7940	27 /100 /704	39 /231 /2523	90 /475 /4652
RedEdge-M_5.5_1280x960 (Green)		233/1400/28163	18 /77 /894	60 /338 /8965	138 /811 /14650
RedEdge-M_5.5_1280x960 (Red)			768/7600/29411	11 /36 /417	17 /62 /524
RedEdge-M_5.5_1280x960 (NIR)				444/4098/28370	260/2016/13454
RedEdge-M_5.5_1280x960 (Red edge)					288/2421/20255

3D Points from 2D Keypoint Matches

	Number of 3D Points Observed
In 2 Images	14347545
In 3 Images	4122336
In 4 Images	1828156
In 5 Images	939852
In 6 Images	574105
In 7 Images	372582
In 8 Images	258187
In 9 Images	186409
In 10 Images	138623
In 11 Images	105258
In 12 Images	81705
In 13 Images	64578
In 14 Images	52054
In 15 Images	42558
In 16 Images	35077
In 17 Images	28519
In 18 Images	23311
In 19 Images	19251
In 20 Images	16232
In 21 Images	13055
In 22 Images	10751
In 23 Images	8967
In 24 Images	7329

In 25 Images	5960
In 26 Images	4948
In 27 Images	4083
In 28 Images	3251
In 29 Images	2556
In 30 Images	2128
In 31 Images	1692
In 32 Images	1347
In 33 Images	1107
In 34 Images	926
In 35 Images	722
In 36 Images	620
In 37 Images	487
In 38 Images	388
In 39 Images	351
In 40 Images	312
In 41 Images	226
In 42 Images	210
In 43 Images	189
In 44 Images	153
In 45 Images	136
In 46 Images	95
In 47 Images	93
In 48 Images	76
In 49 Images	67
In 50 Images	54
In 51 Images	48
In 52 Images	34
In 53 Images	30
In 54 Images	27
In 55 Images	20
In 56 Images	13
In 57 Images	13
In 58 Images	16
In 59 Images	7
In 60 Images	6
In 61 Images	2
In 62 Images	5
In 63 Images	2
In 64 Images	2
In 66 Images	3
In 67 Images	1
In 70 Images	1

2D Keypoint Matches

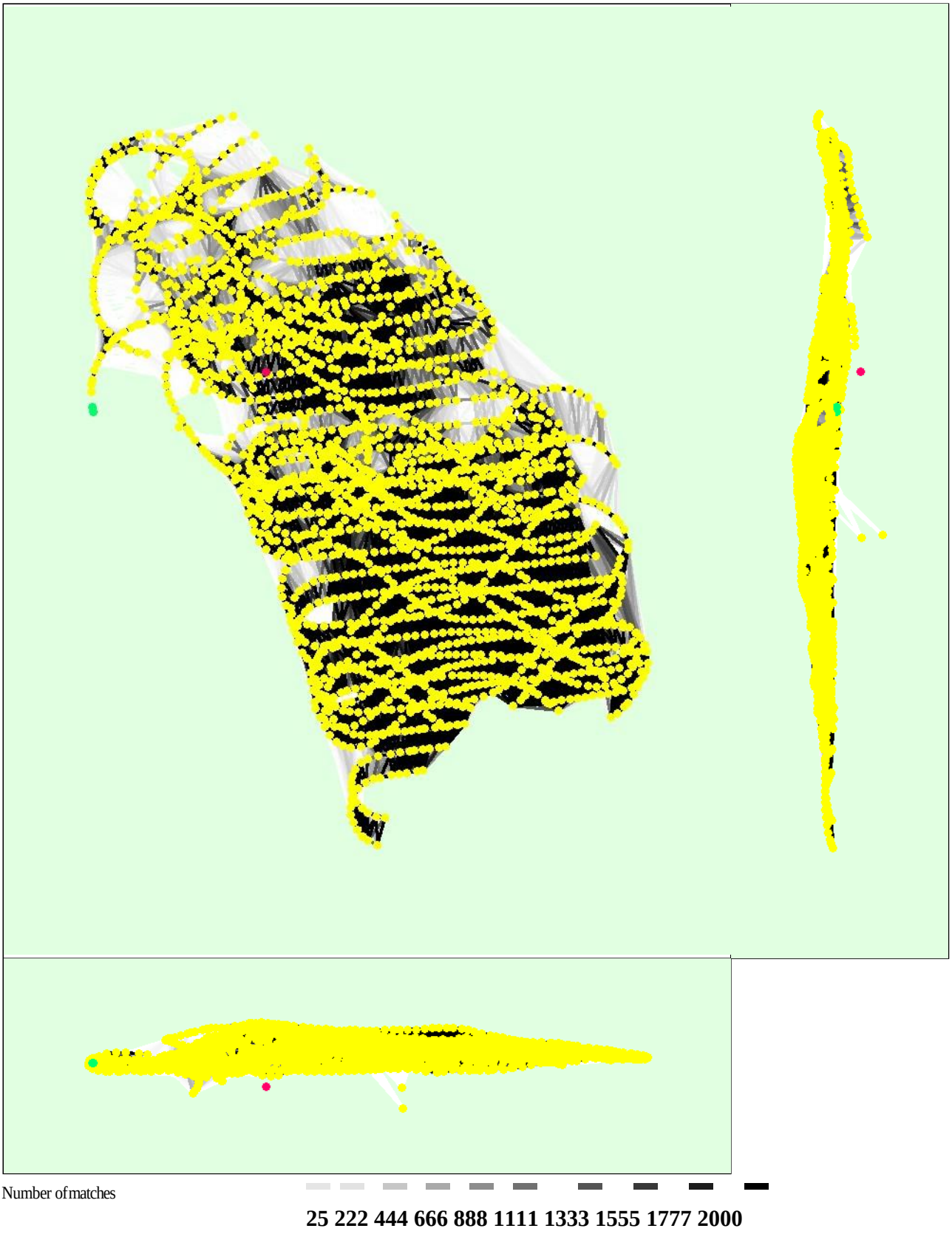


Figure 5: Computed image positions with links between matched images. The darkness of the links indicates the number of matched 2D keypoints between the images. Bright links indicate weak links and require manual tie points or more images.

Geolocation Details

Absolute Geolocation Variance

Min Error [m]	Max Error [m]	Geolocation Error X[%]	Geolocation Error Y[%]	Geolocation Error Z [%]
-	-15.00	0.08	0.00	0.00
-15.00	-12.00	0.13	0.00	0.00
-12.00	-9.00	1.20	0.67	0.00
-9.00	-6.00	5.38	2.03	0.00
-6.00	-3.00	25.23	10.12	0.37
-3.00	0.00	19.14	41.24	50.11
0.00	3.00	17.48	31.38	49.39
3.00	6.00	21.93	11.54	0.13
6.00	9.00	7.35	2.18	0.00
9.00	12.00	1.86	0.76	0.00
12.00	15.00	0.21	0.08	0.00
15.00	-	0.00	0.00	0.00
Mean [m]		0.000566	0.000589	-0.001338
Sigma [m]		4.571686	3.010722	0.918932
RMS Error [m]		4.571686	3.010722	0.918933

Min Error and Max Error represent geolocation error intervals between -1.5 and 1.5 times the maximum accuracy of all the images. Columns X, Y, Z show the percentage of images with geolocation errors within the predefined error intervals. The geolocation error is the difference between the initial and computed image positions. Note that the image geolocation errors do not correspond to the accuracy of the observed 3D points.

Relative Geolocation Variance

Relative Geolocation Error	Images X[%]	Images Y[%]	Images Z [%]
[-1.00, 1.00]	72.43	90.12	100.00
[-2.00, 2.00]	98.05	99.45	100.00
[-3.00, 3.00]	99.92	100.00	100.00
Mean of Geolocation Accuracy [m]	5.000000	5.000000	10.000000
Sigma of Geolocation Accuracy [m]	0.000000	0.000000	0.000000

Images X, Y, Z represent the percentage of images with a relative geolocation error in X, Y, Z.

Initial Processing Details

System Information

Hardware	CPU: Intel(R) Core(TM) i9-9980XE CPU @ 3.00GHz RAM: 128GB GPU: NVIDIA Quadro RTX 4000 (Driver: unknown), NVIDIA Quadro RTX 4000 (Driver: unknown), NVIDIA Quadro RTX 4000 (Driver: unknown), NVIDIA Quadro RTX 4000 (Driver: unknown), NVIDIA Quadro RTX 4000 (Driver: unknown), NVIDIA Quadro RTX 4000 (Driver: unknown), NVIDIA Quadro RTX 4000 (Driver: unknown), NVIDIA Quadro RTX 4000 (Driver: unknown), NVIDIA Quadro RTX 4000 (Driver: unknown), NVIDIA Quadro RTX 4000 (Driver: unknown), NVIDIA Quadro RTX 4000 (Driver: unknown), NVIDIA Quadro RTX 4000 (Driver: unknown), NVIDIA Quadro RTX 4000 (Driver: unknown), NVIDIA Quadro RTX 4000 (Driver: unknown)
Operating System	Windows 10 Education, 64-bit

Coordinate Systems

Image Coordinate System	WGS 84 (EGM96 Geoid)
Output Coordinate System	WGS 84 /UTMzone 30N (EGM96 Geoid)

Processing Options

Detected Template	No Template Available
Keypoints Image Scale	Full, Image Scale: 2
Advanced: Matching Image Pairs	Aerial Grid or Corridor

Advanced: Matching Strategy	Use Geometrically Verified Matching: no
Advanced: Keypoint Extraction	Targeted Number of Keypoints: Automatic
Advanced: Calibration	Calibration Method: Standard Internal Parameters Optimization: All External Parameters Optimization: All Rematch: Custom, yes
Rig «RedEdge-M» processing	optimize relative rotation using a subset of secondarycameras

Point Cloud Densification details

Processing Options

Image Scale	multiscale, 1/2 (Half image size, Default)
Point Density	Optimal
Minimum Number of Matches	3
3D Textured Mesh Generation	yes
3D Textured Mesh Settings:	Resolution: Medium Resolution (default) Color Balancing: no
LOD	Generated: no
Advanced: 3D Textured Mesh Settings	Sample Density Divider: 1
Advanced: Image Groups	Blue, Green, Red, NIR, Red edge
Advanced: Use Processing Area	yes
Advanced: Use Annotations	yes
Time for Point Cloud Densification	15m:59s
Time for Point Cloud Classification	01m:06s
Time for 3D Textured Mesh Generation	16m:55s

Results

Number of Generated Tiles	1
Number of 3D Densified Points	19127920
Average Density(per m^3)	8.71

DSM, Orthomosaic and Index Details

Processing Options

DSMand Orthomosaic Resolution	1 x GSD (7.18 [cm/pixel])
DSMFilters	Noise Filtering: yes Surface Smoothing: yes, Type: Medium
Raster DSM	Generated: yes Method: Inverse Distance Weighting Merge Tiles: yes
Orthomosaic	Generated: yes Merge Tiles: yes GeoTIFF Without Transparency: yes Google Maps Tiles and KML: no
Raster DTM	Generated: yes Merge Tiles: yes
DTMResolution	5 x GSD (7.18 [cm/pixel])
Radiometric calibration with reflectance target	yes
Index Calculator: Reflectance Map	Generated: yes Resolution: 1 x GSD (7.18 [cm/pixel]) Merge Tiles: yes
Index Calculator: Indices	ndvi
Time for DSMGeneration	02m:43s
Time for Orthomosaic Generation	01h:09m:02s
Time for DTMGeneration	03m:26s

Time for Contour Lines Generation	00s
Time for Reflectance Map Generation	01h:43m:54s
Time for Index Map Generation	03m:33s

Camera Radiometric Correction

Camera Name	Band	Radiometric Correction Type	Reflectance target
RedEdge-M_5.5_1280x960	Blue	Camera, Sun Irradiance and Sun Angle using DLS IMU	⚠
RedEdge-M_5.5_1280x960	Green	Camera, Sun Irradiance and Sun Angle using DLS IMU	✓
RedEdge-M_5.5_1280x960	Red	Camera, Sun Irradiance and Sun Angle using DLS IMU	✓
RedEdge-M_5.5_1280x960	NIR	Camera, Sun Irradiance and Sun Angle using DLS IMU	✓
RedEdge-M_5.5_1280x960	Red edge	Camera, Sun Irradiance and Sun Angle using DLS IMU	✓

A.2 Penglaneinon stitching report

Summary

Project	Penglaneinon
Processed	2020-04-17 01:02:00
Camera Model Name(s)	RedEdge_5.5_1280x960 (Blue), RedEdge_5.5_1280x960 (Green), RedEdge_5.5_1280x960 (Red), RedEdge_5.5_1280x960 (NIR), RedEdge_5.5_1280x960 (Red edge)
Rig name(s)	«MicaSense 5 band»
Average Ground Sampling Distance (GSD)	8.73 cm /3.44 in
Area Covered	2.009 km^2 /200.8920 ha /0.78 sq. mi. /496.6719 acres
Time for Initial Processing (without report)	4d:14h:22m:01s

Quality Check

Images	median of 1072 keypoints per image	⚠
Dataset	14427 out of 17390 images calibrated (82%), all images enabled, 2 blocks	⚠
Camera Optimization	0.05% relative difference between initial and optimised internal camera parameters	✓
Matching	median of 684.099 matches per calibrated image	⚠
Georeferencing	yes, no 3D GCP	⚠

Preview

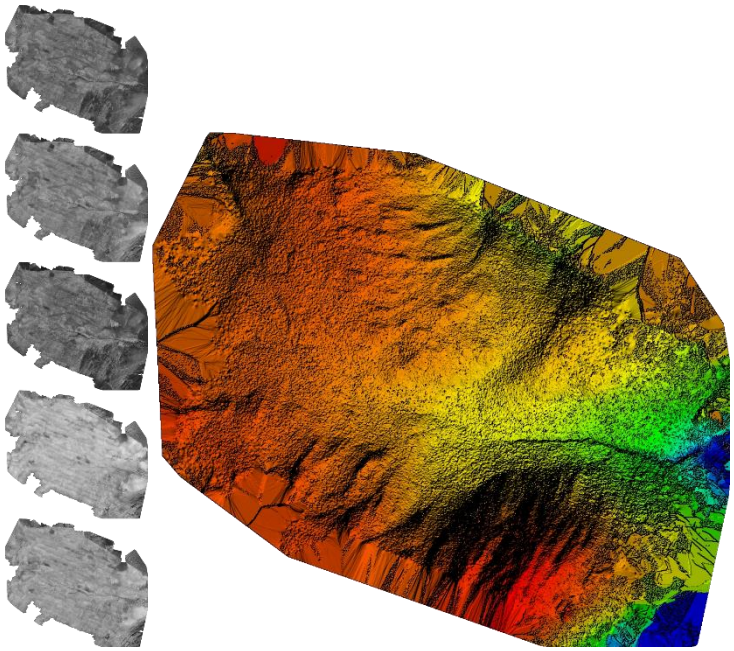


Figure 1: Orthomosaic and the corresponding sparse Digital Surface Model (DSM) before densification.

Calibration Details

Number of Calibrated Images	14427 out of 17390
Number of Geolocated Images	17390 out of 17390

Initial Image Positions

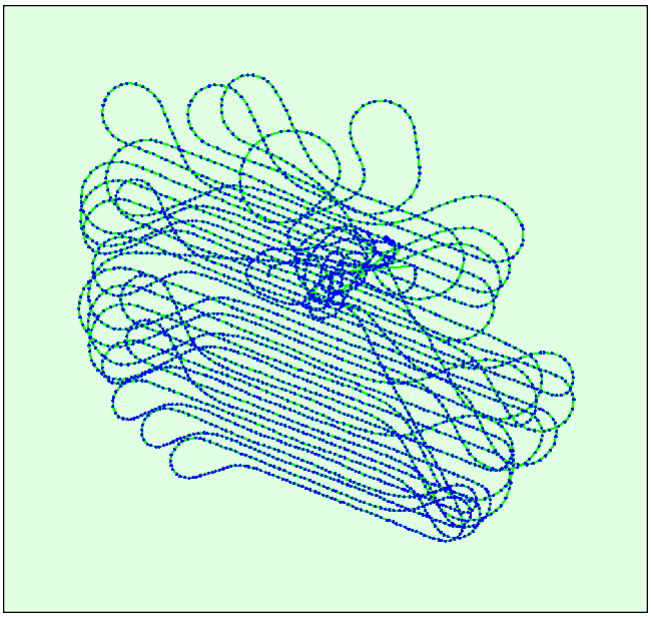
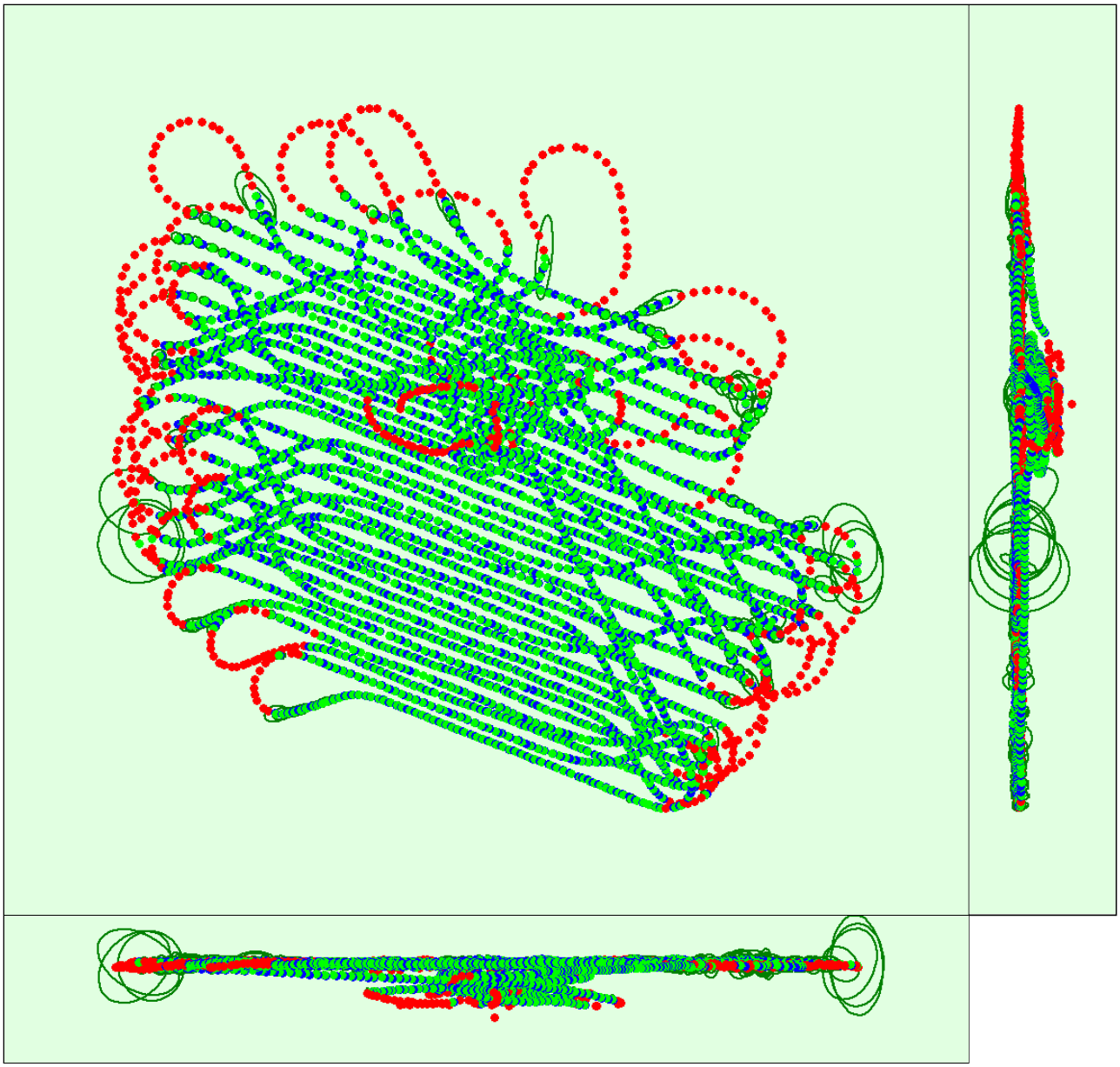


Figure 2: Top view of the initial image position. The green line follows the position of the images in time starting from the large blue dot.

Computed Image/GCPs/Manual Tie Points Positions



Uncertainty ellipses 50xmagnified

Figure 3: Offset between initial (blue dots) and computed (green dots) image positions as well as the offset between the GCPs initial positions (blue crosses) and their computed positions (green crosses) in the top-view (XY plane), front-view (XZ plane), and side-view (YZ plane). Red dots indicate disabled or uncalibrated images. Dark green ellipses indicate the absolute position uncertainty of the bundle block adjustment result.

Absolute camera position and orientation uncertainties

	X[m]	Y[m]	Z [m]	Omega [degree]	Phi [degree]	Kappa [degree]
Mean	0.130	0.120	0.162	0.056	0.060	0.022
Sigma	0.077	0.088	0.083	0.088	0.072	0.067

Overlap

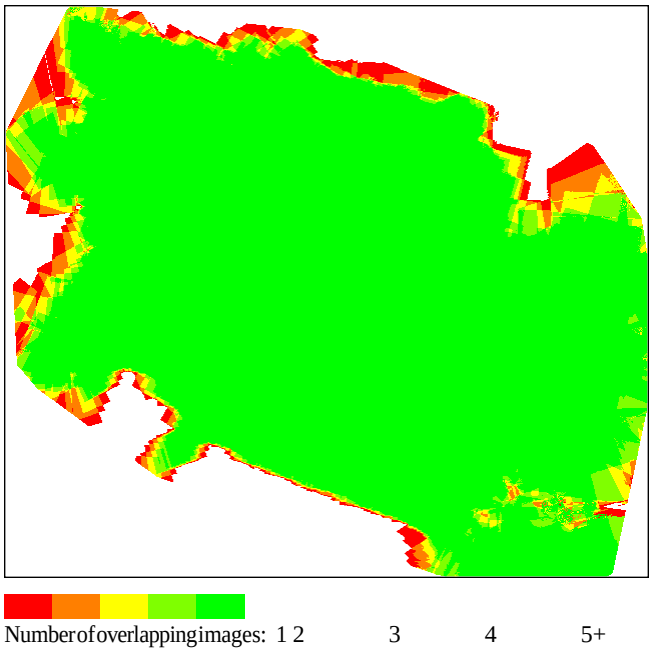


Figure 4: Number of overlapping images computed for each pixel of the orthomosaic.
Red and yellow areas indicate low overlap for which poor results may be generated. Green areas indicate an overlap of over 5 images for every pixel. Good quality results will be generated as long as the number of keypoint matches is also sufficient for these areas (see Figure 5 for keypoint matches).

Bundle Block Adjustment Details

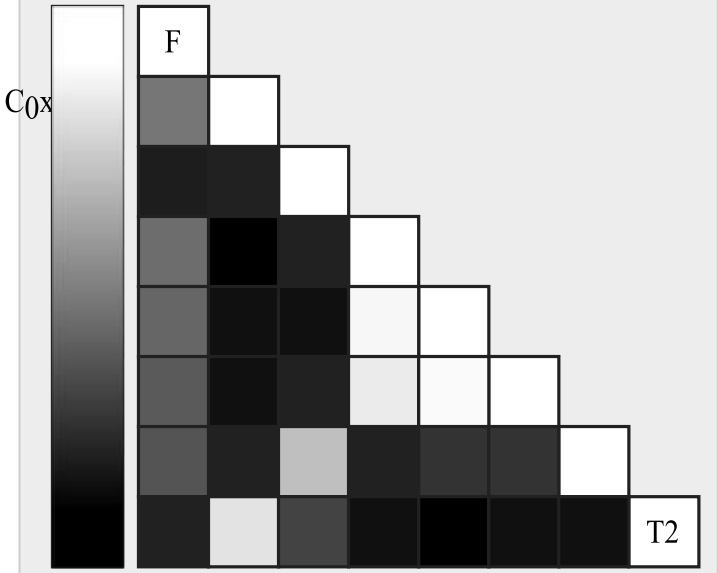
Number of 2D Keypoint Observations for Bundle Block Adjustment	2730725
Number of 3D Points for Bundle Block Adjustment	568194
Mean Reprojection Error [pixels]	0.145

Internal Camera Parameters

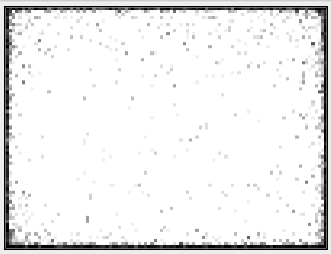
RedEdge_5.5_1280x960 (Blue). Sensor Dimensions: 4.800 [mm] x 3.600 [mm]

EXIF ID: RedEdge_5.5_1280x960

	Focal Length	Principal Point x	Principal Point y	R1	R2	R3	T1	T2
Initial Values	1440.273 [pixel] 5.401 [mm]	632.645 [pixel] 2.372 [mm]	467.285 [pixel] 1.752 [mm]	-0.103	0.166	-0.055	-0.000	0.001
Optimised Values	1439.548 [pixel] 5.398 [mm]	630.726 [pixel] 2.365 [mm]	470.292 [pixel] 1.764 [mm]	-0.103	0.169	-0.077	-0.000	0.000
Uncertainties (Sigma)	0.201 [pixel] 0.001 [mm]	0.232 [pixel] 0.001 [mm]	0.180 [pixel] 0.001 [mm]	0.002	0.012	0.026	0.000	0.000



The correlation between camera internal parameters determined by the bundle adjustment. White indicates a full correlation between the parameters, ie. any change in one can be fully compensated by the other. Black indicates that the parameter is completely independent, and is not affected by other parameters.

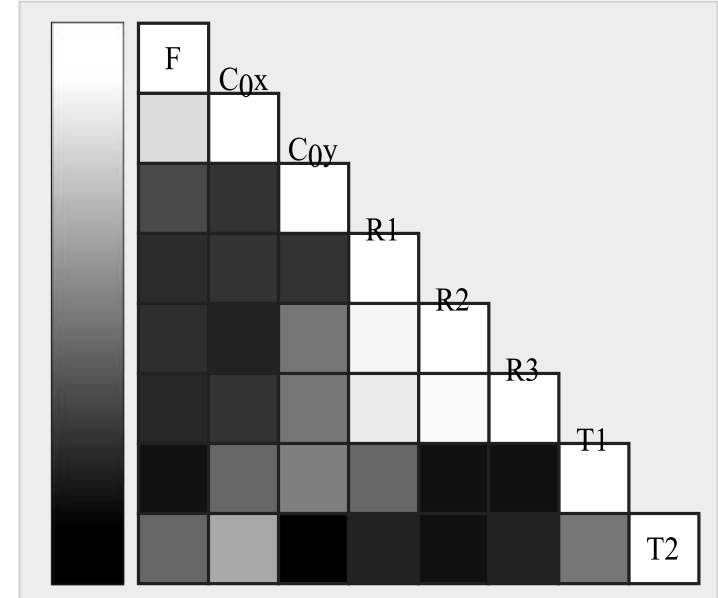


The number of Automatic Tie Points (ATPs) per pixel, averaged over all images of the camera model, is color coded between black and white. White indicates that, on average, more than 16 ATPs have been extracted at the pixel location. Black indicates that, on average, 0 ATPs have been extracted at the pixel location. Click on the image to the see the average direction and magnitude of the re- projection error for each pixel. Note that the vectors are scaled for better visualization. The scale bar indicates the magnitude of 1 pixel error.

Internal Camera Parameters

RedEdge_5.5_1280x960 (Green). Sensor Dimensions: 4.800 [mm] x 3.600 [mm]
EXIF ID: RedEdge_5.5_1280x960

	Focal Length	Principal Pointx	Principal Pointy	R1	R2	R3	T1	T2
Initial Values	1448.621 [pixel] 5.432 [mm]	654.443 [pixel] 2.454[mm]	484.776 [pixel] 1.818[mm]	-0.103	0.176	-0.087	-0.000	0.000
Optimized Values	1448.034 [pixel] 5.430 [mm]	652.387[pixel] 2.446[mm]	488.970 [pixel] 1.834[mm]	-0.099 0.001	0.156 0.004	-0.050 0.009	0.000 0.000	-0.000 0.000



The correlation between camera internal parameters determined by the bundle adjustment. White indicates a full correlation between the parameters, ie. any change in one can be fully compensated by the other. Black indicates that the parameter is completely independent, and is not affected by other parameters.

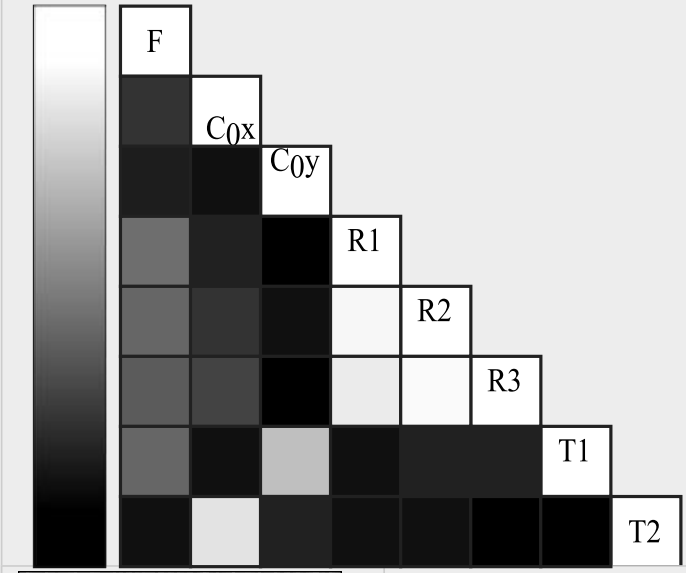
The number of Automatic Tie Points (ATPs) per pixel, averaged over all images of the camera model, is color coded between black and white. White indicates that, on average, more than 16 ATPs have been extracted at the pixel location. Black indicates that, on average, 0 ATPs have been extracted at the pixel location. Click on the image to the see the average direction and magnitude of the re- projection error for each pixel. Note that the vectors are scaled for better visualization. The scale bar indicates the magnitude of 1 pixel error.

Internal Camera Parameters

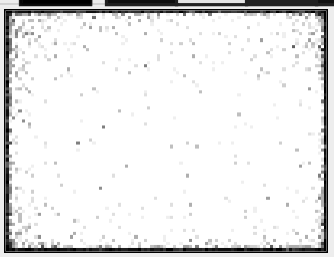
RedEdge_5.5_1280x960 (Red). Sensor Dimensions: 4.800 [mm] x 3.600 [mm]

EXIF ID: RedEdge_5.5_1280x960

	Focal Length	Principal Point x	Principal Point y	R1	R2	R3	T1	T2
Initial Values	1456.812 [pixel] 5.463 [mm]	655.973 [pixel] 2.460 [mm]	479.352 [pixel] 1.798 [mm]	-0.107	0.177	-0.087	-0.000	0.001
Optimised Values	1455.903 [pixel] 5.460 [mm]	656.737 [pixel] 2.463 [mm]	481.202 [pixel] 1.805 [mm]	-0.105	0.169	-0.070	-0.000	0.000
Uncertainties (Sigma)	0.203 [pixel] 0.001 [mm]	0.240 [pixel] 0.001 [mm]	0.186 [pixel] 0.001 [mm]	0.002	0.012	0.027	0.000	0.000



The correlation between camera internal parameters determined by the bundle adjustment. White indicates a full correlation between the parameters, ie. any change in one can be fully compensated by the other. Black indicates that the parameter is completely independent, and is not affected by other parameters.



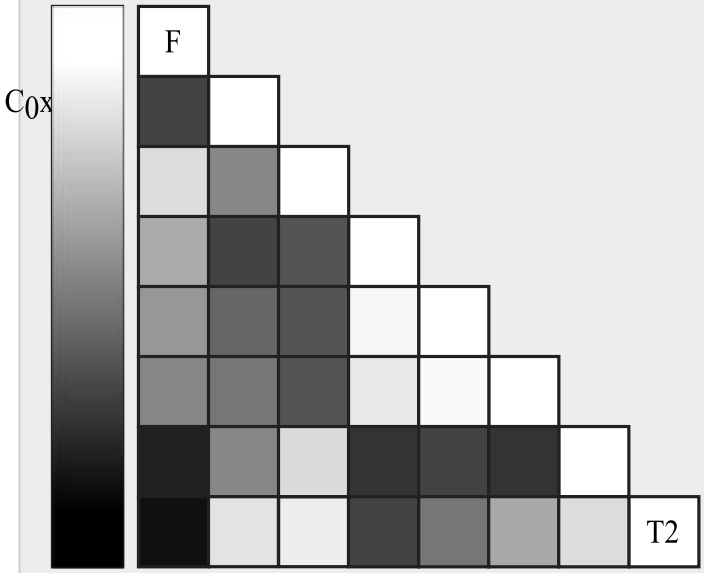
The number of Automatic Tie Points (ATPs) per pixel, averaged over all images of the camera model, is color coded between black and white. White indicates that, on average, more than 16 ATPs have been extracted at the pixel location. Black indicates that, on average, 0 ATPs have been extracted at the pixel location. Click on the image to the see the average direction and magnitude of the re- projection error for each pixel. Note that the vectors are scaled for better visualization. The scale bar indicates the magnitude of 1 pixel error.

Internal Camera Parameters

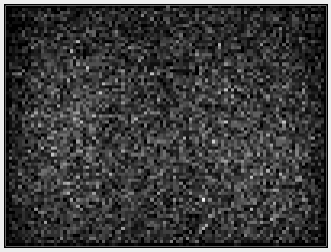
RedEdge_5.5_1280x960 (NIR). Sensor Dimensions: 4.800 [mm] x 3.600 [mm]

EXIF ID: RedEdge_5.5_1280x960

	Focal Length	Principal Point x	Principal Point y	R1	R2	R3	T1	T2
Initial Values	1458.205 [pixel] 5.468 [mm]	671.299 [pixel] 2.517 [mm]	489.963 [pixel] 1.837 [mm]	-0.109	0.184	-0.109	-0.000	0.000
Optimised Values	1457.023 [pixel] 5.464 [mm]	672.328 [pixel] 2.521 [mm]	491.542 [pixel] 1.843 [mm]	-0.105	0.152	-0.045	-0.000	0.000
Uncertainties (Sigma)	0.274 [pixel] 0.001 [mm]	0.540 [pixel] 0.002 [mm]	0.462 [pixel] 0.002 [mm]	0.003	0.024	0.053	0.000	0.000



The correlation between camera internal parameters determined by the bundle adjustment. White indicates a full correlation between the parameters, ie. any change in one can be fully compensated by the other. Black indicates that the parameter is completely independent, and is not affected by other parameters.

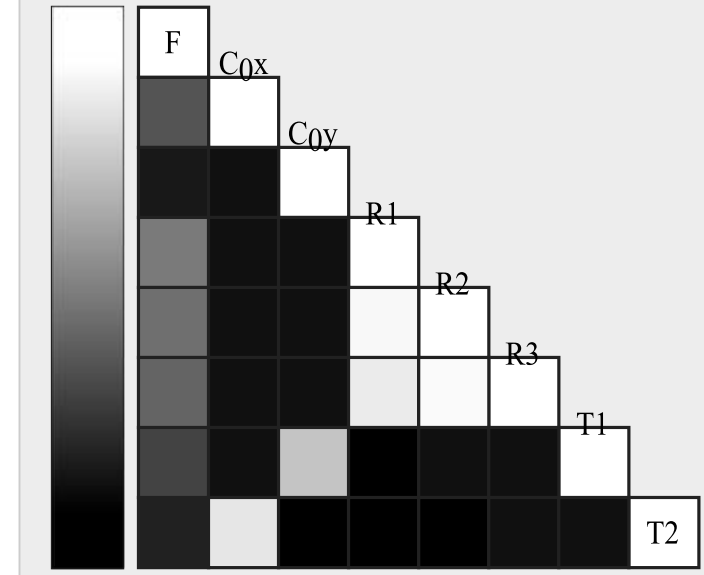


The number of Automatic Tie Points (ATPs) per pixel, averaged over all images of the camera model, is color coded between black and white. White indicates that, on average, more than 16 ATPs have been extracted at the pixel location. Black indicates that, on average, 0 ATPs have been extracted at the pixel location. Click on the image to the see the average direction and magnitude of the re- projection error for each pixel. Note that the vectors are scaled for better visualization. The scale bar indicates the magnitude of 1 pixel error.

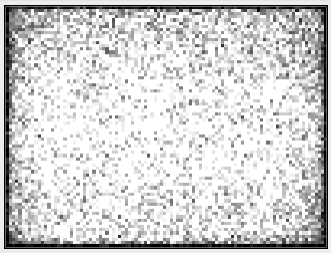
Internal Camera Parameters

RedEdge_5.5_1280x960 (Red edge). Sensor Dimensions: 4.800 [mm] x 3.600 [mm]
EXIF ID: RedEdge_5.5_1280x960

	Focal Length	Principal Point x	Principal Point y	R1	R2	R3	T1	T2
Initial Values	1449.139 [pixel] 5.434 [mm]	639.629 [pixel] 2.399 [mm]	472.899 [pixel] 1.773 [mm]	-0.103	0.147	-0.026	0.000	0.000
Optimized Values	1448.512 [pixel] 5.432 [mm]	640.446 [pixel] 2.402 [mm]	476.016 [pixel] 1.785 [mm]	-0.104	0.154	-0.042	0.000	-0.000
				0.002	0.014	0.033	0.000	0.000



The correlation between camera internal parameters determined by the bundle adjustment. White indicates a full correlation between the parameters, ie. any change in one can be fully compensated by the other. Black indicates that the parameter is completely independent, and is not affected by other parameters.



The number of Automatic Tie Points (ATPs) per pixel, averaged over all images of the camera model, is color coded between black and white. White indicates that, on average, more than 16 ATPs have been extracted at the pixel location. Black indicates that, on average, 0 ATPs have been extracted at the pixel location. Click on the image to the see the averaged direction and magnitude of there- projection error for each pixel. Note that the vectors are scaled for better visualization. The scale bar indicates the magnitude of 1 pixel error.

Camera Rig «MicaSense 5 band» Relatives. Images: 17390

	Transl X[m]	Transl Y[m]	TranslZ[m]	Rot X[degree]	Rot Y[degree]	RotZ[degree]
RedEdge_5.5_1280x960 (Green)	Reference Camera					
RedEdge_5.5_1280x960 (Blue)						
Initial Values	0.030	0.000	0.000	0.000	0.000	0.000
Optimised values	0.030	0.000	0.000	-0.097	-0.293	-0.333
Uncertainties (sigma)				0.007	0.010	0.001
RedEdge_5.5_1280x960 (Red)						
Initial Values	0.000	0.022	0.000	0.000	0.000	0.000
Optimised values	0.000	0.022	0.000	-0.062	0.165	-0.314
Uncertainties (sigma)				0.007	0.010	0.001
RedEdge_5.5_1280x960 (NIR)						
Initial Values	0.030	0.022	0.000	0.000	0.000	0.000
Optimised values	0.030	0.022	0.000	0.066	-0.287	-0.381
Uncertainties (sigma)				0.019	0.022	0.002
RedEdge_5.5_1280x960 (Red edge)						
Initial Values	0.015	0.011	0.000	0.000	0.000	0.000
Optimised values	0.015	0.011	0.000	0.043	-0.198	-0.288
Uncertainties (sigma)				0.008	0.011	0.001

2D Keypoints Table

	Number of 2D Keypoints per Image	Number of Matched 2D Keypoints per Image
Median	1072	684
Min	434	0
Max	2241	1309
Mean	1070	669

2D Keypoints Table for Camera RedEdge_5.5_1280x960 (Blue)

	Number of 2D Keypoints per Image	Number of Matched 2D Keypoints per Image
Median	1189	723
Min	781	62
Max	1718	1309
Mean	1185	701

2D Keypoints Table for Camera RedEdge_5.5_1280x960 (Green)

	Number of 2D Keypoints per Image	Number of Matched 2D Keypoints per Image
Median	1073	714
Min	524	71
Max	2241	1207
Mean	1072	712

2D Keypoints Table for Camera RedEdge_5.5_1280x960 (Red)

	Number of 2D Keypoints per Image	Number of Matched 2D Keypoints per Image
Median	1248	752
Min	818	25
Max	1698	1241
Mean	1239	730

2D Keypoints Table for Camera RedEdge_5.5_1280x960 (NIR)

	Number of 2D Keypoints per Image	Number of Matched 2D Keypoints per Image
Median	826	302
Min	434	0
Max	1905	792
Mean	844	315

2D Keypoints Table for Camera RedEdge_5.5_1280x960 (Red edge)

	Number of 2D Keypoints per Image	Number of Matched 2D Keypoints per Image
Median	992	506
Min	577	73
Max	1834	864
Mean	991	512

Median /75%/Maximal Number of Matches Between Camera Models

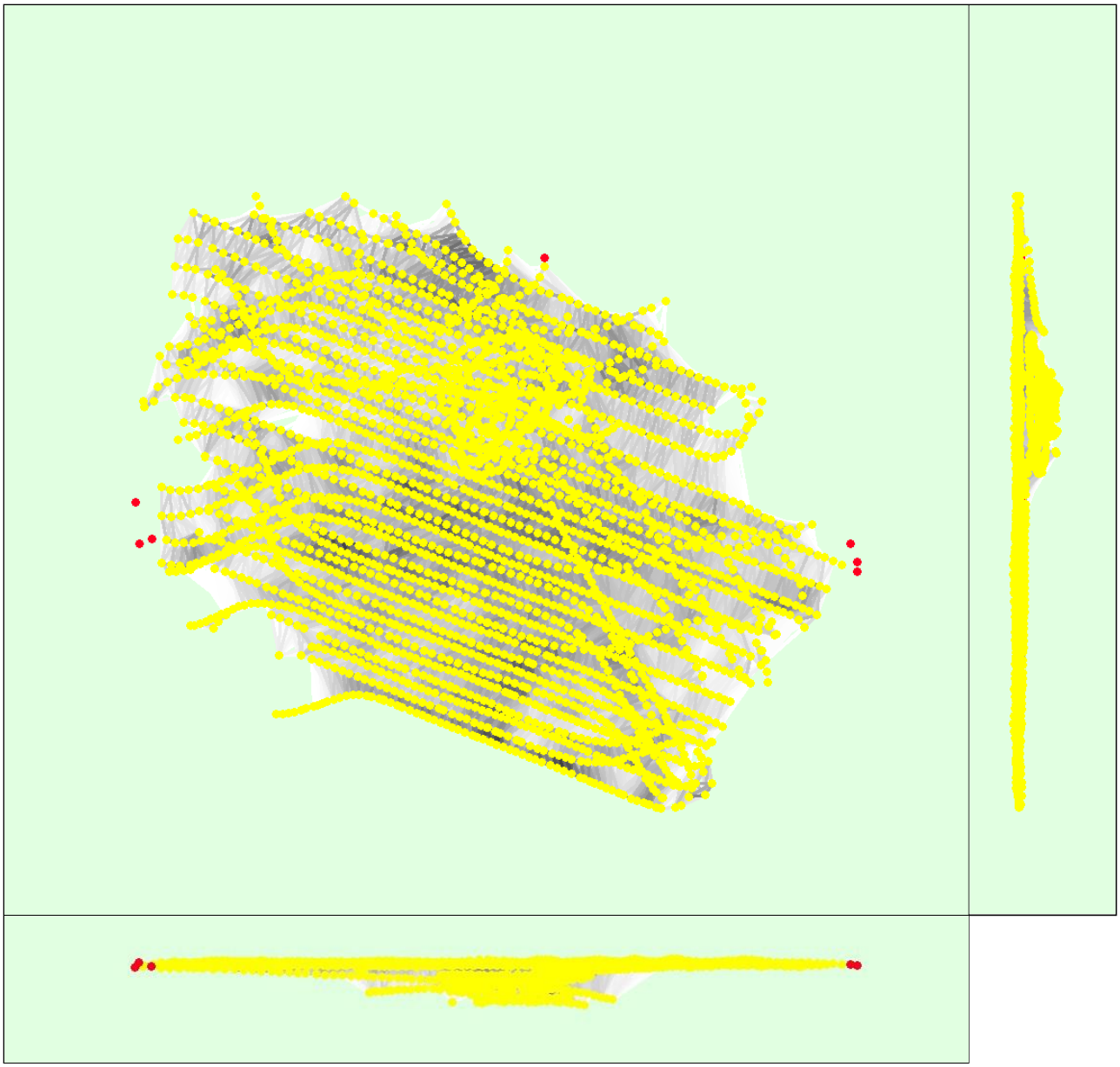
	RedEdge_5.5_12... (Blue)	RedEdge_5.5_1... (Green)	RedEdge_5.5_128... (Red)	RedEdge_5.5_128... (NIR)	RedEdge_5..... (Red edge)
RedEdge_5.5_1280x960 (Blue)	57 /166 /787	37 /99 /726	65 /183 /812	2 /4 /24	35 /89 /437
RedEdge_5.5_1280x960 (Green)		39 /108 /738	36 /97 /694	2 /4 /54	26 /70 /519
RedEdge_5.5_1280x960 (Red)			61 /182 /845	2 /4 /31	35 /89 /436
RedEdge_5.5_1280x960 (NIR)				59 /186 /643	3 /7 /113
RedEdge_5.5_1280x960 (Red edge)					37 /117 /547

3D Points from 2D Keypoint Matches

	Number of 3D Points Observed
In 2 Images	270008
In 3 Images	86848
In 4 Images	48354
In 5 Images	30250
In 6 Images	22195
In 7 Images	17028
In 8 Images	13729
In 9 Images	11229
In 10 Images	9304
In 11 Images	7840
In 12 Images	6797
In 13 Images	5965
In 14 Images	5035
In 15 Images	4387
In 16 Images	3773
In 17 Images	3378
In 18 Images	3041
In 19 Images	2742
In 20 Images	2245

In 21 Images	1947
In 22 Images	1664
In 23 Images	1470
In 24 Images	1346
In 25 Images	1145
In 26 Images	961
In 27 Images	836
In 28 Images	680
In 29 Images	602
In 30 Images	496
In 31 Images	417
In 32 Images	335
In 33 Images	296
In 34 Images	257
In 35 Images	197
In 36 Images	212
In 37 Images	197
In 38 Images	143
In 39 Images	130
In 40 Images	124
In 41 Images	87
In 42 Images	81
In 43 Images	65
In 44 Images	68
In 45 Images	43
In 46 Images	40
In 47 Images	28
In 48 Images	36
In 49 Images	27
In 50 Images	16
In 51 Images	12
In 52 Images	13
In 53 Images	8
In 54 Images	10
In 55 Images	5
In 56 Images	2
In 57 Images	8
In 58 Images	9
In 59 Images	7
In 60 Images	4
In 61 Images	2
In 62 Images	4
In 63 Images	3
In 65 Images	3
In 66 Images	2
In 67 Images	1
In 68 Images	1
In 69 Images	1
In 70 Images	1
In 72 Images	1
In 73 Images	1
In 74 Images	1
In 76 Images	1

2D Keypoint Matches



Number of matches

25 90 180 270 360 451 541 631 721 812

Figure 5: Computed image positions with links between matched images. The darkness of the links indicates the number of matched 2D keypoints between the images. Bright links indicate weak links and require manual tie points or more images.

Geolocation Details

Absolute Geolocation Variance

Min Error [m]	Max Error [m]	Geolocation Error X[%]	Geolocation Error Y[%]	Geolocation Error Z [%]
-	-15.00	0.03	0.07	0.00
-15.00	-12.00	0.21	0.20	0.00
-12.00	-9.00	1.42	0.57	0.00
-9.00	-6.00	8.16	1.80	0.00
-6.00	-3.00	28.87	10.59	0.35
-3.00	0.00	15.47	36.31	45.08
0.00	3.00	13.29	38.90	54.36
3.00	6.00	19.19	10.27	0.21
6.00	9.00	9.11	1.09	0.00
9.00	12.00	2.96	0.07	0.00
12.00	15.00	0.87	0.10	0.00
15.00	-	0.42	0.03	0.00
Mean [m]		0.000433	0.002386	-0.000419
Sigma [m]		5.211401	2.976490	1.013199
RMS Error [m]		5.211402	2.976491	1.013199

Min Error and Max Error represent geolocation error intervals between -1.5 and 1.5 times the maximum accuracy of all the images. Columns X, Y, Z show the percentage of images with geolocation errors within the predefined error intervals. The geolocation error is the difference between the initial and computed image positions. Note that the image geolocation errors do not correspond to the accuracy of the observed 3D points.

Relative Geolocation Variance

Relative Geolocation Error	Images X[%]	Images Y[%]	Images Z [%]
[-1.00, 1.00]	65.89	93.62	100.00
[-2.00, 2.00]	96.11	99.17	100.00
[-3.00, 3.00]	99.55	99.90	100.00
Mean of Geolocation Accuracy [m]	5.000000	5.000000	10.000000
Sigma of Geolocation Accuracy [m]	0.000000	0.000000	0.000000

Images X, Y, Z represent the percentage of images with a relative geolocation error in X, Y, Z.

Initial Processing Details

System Information

Hardware	CPU: Intel(R) Core(TM) i9-9980XE CPU @ 3.00GHz RAM: 128GB GPU: NVIDIA Quadro RTX 4000 (Driver: unknown), NVIDIA Quadro RTX 4000 (Driver: unknown), NVIDIA Quadro RTX 4000 (Driver: unknown), NVIDIA Quadro RTX 4000 (Driver: unknown), NVIDIA Quadro RTX 4000 (Driver: unknown), NVIDIA Quadro RTX 4000 (Driver: unknown), NVIDIA Quadro RTX 4000 (Driver: unknown), NVIDIA Quadro RTX 4000 (Driver: unknown), NVIDIA Quadro RTX 4000 (Driver: unknown), NVIDIA Quadro RTX 4000 (Driver: unknown), NVIDIA Quadro RTX 4000 (Driver: unknown), NVIDIA Quadro RTX 4000 (Driver: unknown), NVIDIA Quadro RTX 4000 (Driver: unknown), NVIDIA Quadro RTX 4000 (Driver: unknown)
Operating System	Windows 10 Education, 64-bit

Coordinate Systems

Image Coordinate System	WGS 84 (EGM96 Geoid)
Output Coordinate System	WGS 84 /UTMzone 30N (EGM96 Geoid)

Processing Options

Detected Template	No Template Available
Keypoints Image Scale	Custom, Image Scale: 0.5
Advanced: Matching Image Pairs	Aerial Grid or Corridor
Advanced: Matching Strategy	Use Geometrically Verified Matching: yes
Advanced: Keypoint Extraction	Targeted Number of Keypoints: Automatic
Advanced: Calibration	Calibration Method: Alternative Internal Parameters Optimization: All External Parameters Optimization: All Rematch: Custom, yes
Rig «MicaSense 5 band» processing	optimize relative rotation using a subset of secondarycameras

Point Cloud Densification details

Processing Options

Image Scale	multiscale, 1/2 (Half image size, Default)
Point Density	Optimal
Minimum Number of Matches	3
3D Textured Mesh Generation	no
LOD	Generated: no
Advanced: Image Groups	Blue, Green, Red, NIR, Red edge
Advanced: Use Processing Area	yes
Advanced: Use Annotations	yes
Time for Point Cloud Densification	14m:16s
Time for Point Cloud Classification	01m:18s
Time for 3D Textured Mesh Generation	NA

Results

Number of Generated Tiles	1
Number of 3D Densified Points	17309391
Average Density(per m^3)	5.46

DSM, Orthomosaic and Index Details

Processing Options

DSMand Orthomosaic Resolution	1 x GSD (8.73 [cm/pixel])
DSMFilters	Noise Filtering: yes Surface Smoothing: yes, Type: Sharp
Raster DSM	Generated: yes Method: Inverse Distance Weighting Merge Tiles: yes
Orthomosaic	Generated: yes Merge Tiles: yes GeoTIFF Without Transparency: yes Google Maps Tiles and KML: no
Raster DTM	Generated: yes Merge Tiles: yes
DTMResolution	5 x GSD (8.73 [cm/pixel])
Radiometric calibration with reflectance target	yes
Index Calculator: Reflectance Map	Generated: yes Resolution: 1 x GSD (8.73 [cm/pixel]) Merge Tiles: yes
Index Calculator: Indices	ndvi
Index Calculator: Index Values	Polygon Shapefile [cm/grid]: 400
Time for DSMGeneration	04m:51s
Time for Orthomosaic Generation	42m:38s
Time for DTMGeneration	02m:55s
Time for Contour Lines Generation	00s
Time for Reflectance Map Generation	01h:20m:58s
Time for Index Map Generation	03m:50s

Camera Radiometric Correction

Camera Name	Band	Radiometric Correction Type	Reflectance target
RedEdge_5.5_1280x960	Blue	Camera and Sun Irradiance	✓
RedEdge_5.5_1280x960	Green	Camera and Sun Irradiance	✓
RedEdge_5.5_1280x960	Red	Camera and Sun Irradiance	✓
RedEdge_5.5_1280x960	NIR	Camera and Sun Irradiance	✓
RedEdge_5.5_1280x960	Red edge	Camera and Sun Irradiance	✓

A.3 Beferrer parameters

-a,--Activity <b1,s1,b2,s2...>

- Activity category specification based on speed. Provide behaviour name (bn) and maximum speed (sn) for each behaviour required as a comma separated list. No spaces in parameter list. Default is <Stationary,0.02,Foraging,0.33,Moving ,10.0>

- c,--CoordinatesFence <N,W,S,E>

- Ignore coordinates outside the region <N,W,S,E>. Comma separated decimal lat, log ,lat ,lon values. No spaces in parameter list.

- cde,--EndDate <dd/mm/yyyy>

- Ignore data after the specified date. Must use 4 digit year.

- cds,--StartDate <dd/mm/yyyy>

- Ignore data before the specified date. Must use 4 digit year.

- cg,--GlitchSpeed <speed>

- Minimum speed considered as a single point gps glitch. Data point will be removed if adjacent points provide movement with speed below the specified value.

- cte,--EndTime <hh:mm:ss>

- Ignore daily data after the specified time.

- cts,--StartTime <hh:mm:ss>

- Ignore daily data before the specified time. If StartTime greater than EndTime then overnight period.

- f,--File <name>

- Input file name (including path if required). Ignored if - fd used.

- fd,--Folder <folder>

- Input folder/directory containing data files. If neither file (-f) or folder (-fd) is specified all files in the current working folder will be used subject to any suffix specification set with - fs

- fs,--Suffix <suffix>

- Suffix to be used to select input file names. Include the '.' when the suffix is a file type extension. e.g. .TXT Do not use for - f option where complete filename is used. Omission in conjunction with - fd will process every file in the folder.

- g,--GroupTime <min:max>

- Time resolution of data to consider. Each sequential set of coordinates that satisfy this range is considered a sample group i.e. a track segment that includes movements between min and max seconds duration. Extracts sample group sequences from a mixed resolution data set. Default is all data.

- ge,--TrimGroupEnd <n>

- Remove n points from the end of each sample group.

- gg,--TrimGroupBelowSpeed <speed>

- All movements before or after a movement greater than the specified speed are removed from each group.

- gs,--TrimGroupStart <n>

- Remove n of points from the start of each sample group.

- h,--help

- This page. In case of problems or queries contact nns@aber.ac.uk .

- m,--MovementAverage <time>

- Average the movement speed data over the specified time period (seconds). Omission or 0 value will result in no averaging of movements.

- p,--PositionAverage <dist>

- Average the position coordinates over the specified distance (meters) prior to generating movements. Smooths gps track in the presence of gps lockup or gps noise. Alternatively set to 0 to combine non-moving sequences (gps lockup) into a single longer duration move. Default if not set 2.0m

- v,--Verbose

- Detailed output of the processing steps. Use for explanation or debugging (can be long so redirect to a file).

Beferrer accepts comma delimited text files as input. Each valid line should contain four items: lat, long, date, time. Lat and long are decimals. Dates are in dd/mm/yyyy format and times in hh:mm:ss format. An example of a valid line of data is: 52.439037, - 3.993740, 27/5/2018, 10:5:54

Any lines that do not contain four items in this format are ignored.

Output files are created in the same folder as the input data.

These file names include a compact version of the (non-default) command line options to allow output to be easily preserved if the tool is run with different options.

A summary of the processing steps is output to the terminal. More detailed information can be produced using the - v (or - -Verbose) option.

The tool also produces several data files for the widely used (free) gnuplot tool to allow quick visual sanity checking of the results. Obtain gnuplot from: http://www.gnuplot.info

Assuming that gnuplot is properly installed, the contents of the file prefixed 'gnuplot_commands' can be pasted onto the command line (after setting the working folder to the location of the output) to produce .eps graphs.

These files contain gnuplot commands and the relevant data, in plain text format and can be edited as desired, e.g. to change line widths, axis ranges, titles etc.

A.4 Methods in Ecology and Evolution paper (Published)

Handling editor: Dr Theoni Photopoulou

Article type : Practical Tools

First published: 08 August 2020

https://doi.org/10.1111/2041-210X.13464

A bespoke low-cost system for radio tracking animals using multi-rotor and fixed-wing unmanned aerial vehicles (UAVs)

Benjamin Roberts[1*], Mark Neal[2], Neal Snooke[3], Frédéric Labrosse[3], Tristan Curteis[4], and Mariecia Fraser[1]

*Correspondence author. E-mail: ber32@aber.ac.uk

[1] Pwllpeiran Upland Research Centre, Aberystwyth University, Cwmystwyth, Aberystwyth, United Kingdom, SY23 4AB, UK
[2] Ystumtec Ltd, Ystumtuen, Aberystwyth, United Kingdom, SY23 3AF, UK
[3] Department of Computer Science, Aberystwyth University, Penglais Campus, Aberystwyth, United Kingdom, SY23 3FL, UK
[4] Penywern, Llanrhaeadr Ym Mochnant, Oswestry, United Kingdom, SY10 0BN, UK

Abstract

1. Due to the costs of related technologies tracking studies typically use low numbers of animals as representative samples for whole group or species analysis, often without clear knowledge as to how representative these numbers are.
2. The use of unmanned aerial vehicles (UAVs) has the potential to considerably improve radio frequency (RF) based tracking systems. This includes improved line-of-sight visibility, access and range in difficult terrain, and an increase in achievable spatial accuracy.
3. This paper presents details of a fully custom-built active RFID tag and receiver

system bespoke to UAVs, compatible with both multi-rotor and fixed-wing platforms. Using sheep as a model we show the suitability of this system for tracking large terrestrial mammals.

4. During static testing using both platform types we calculated a spatial accuracy of 58.5 m (based on 95th percentile/R95 parameter) for this system using data from 14 flights (n = 175 tag interactions). When tested on sheep, working tags were detected 93% of the time over 7 conducted flights.
5. We provide practical considerations for operating this system on a UAV platform, address concerns relating to the system, and identify future areas of research both for this system and other UAV based RF tracking systems.

Keywords

Applied Ecology, Monitoring, Conservation, Agricultural systems

Introduction

The spatial and temporal distribution of animals is frequently a foundation for understanding biological phenomena within physiological, behavioural and ecological studies (Kays *et al.* 2015). The increased utilisation of GPS in recent years has led to refinement in the achievable accuracy of animal tracking devices, and reductions in the labour required to operate them. However, with this has come an increase in cost which often corresponds to low numbers of animals being tracked, and the assumption that the positional information of a subset of individuals is representative of whole herd/group movements. While RF tags cannot provide the continuous tracking capability of GPS equipped trackers they are inexpensive and can be extremely small and lightweight, allowing large number of animals to be tracked albeit at lower spatial precision and frequency.

Advances in the autonomous capability and payload capacity of unmanned aerial systems has led to them being increasingly utilised and explored as potential data collection platforms in ecological surveying and monitoring (Hodgson *et al.* 2018). The ability of unmanned aerial vehicles (UAVs) to travel long ranges quickly (particularly fixed-wing UAVs) whilst offering greater predictable likelihood of line-of-sight of target animals (Körner *et al.* 2010) provides advantages over conventional methods of radio tracking on the ground.

Recently, researchers have begun to explore the potential benefits of UAV-based radio tracking systems (hereon referred to as UAVRTS) compared to conventional methods. However, as Shafer *et al* (2019) note, many of the presented systems exist primarily as proof-of-principle concepts. The prime focus in most of these studies is the refinement of the localisation methods employed. Whilst this may be valuable in considering potential hardware configuration options, there remains sizeable knowledge gaps within the subject area that have delayed the development of field-ready systems. Firstly, there has been very limited testing on animals, with tagging to date almost exclusively restricted to avian species (Cliff *et al.* 2015; Tremblay *et al.* 2017). Furthermore, many studies are limited to single tag testing (Körner *et al.* 2010; Dos Santos *et al.* 2014; Bayram, Stefas & Isler 2018; Shafer *et al.* 2019), and thus their ability to track movements when multiple animals are tagged remains unknown. Furthermore, none of the studies have utilised or tested their systems on fixed-wing UAVs. Given that fixed-wing UAVs offer vastly superior range, flight speed and endurance compared to multi rotor platforms, there is an opportunity to greatly expand the capability of UAVRTS by using such a platform.

The novel system reported in this paper features a fully custom-made active radio-frequency identification (RFID) tag and receiver system suitable for both fixed-wing and multi-rotor

UAVs. The electronic components are purposely low cost with the goal of making tagging greater numbers of animals more affordable. Unlike previous studies where existing commercial tags have been used or modified, we present a bespoke tag specifically designed for detection by a UAVRTS. Whereas most existing tags continually transmit when activated, our RFID tags remain in a dormant state, with a brief listening period occurring every 6 s. Tag responses are only elicited when a tag exciter trigger located on the UAV comes into operation, thereby saving considerable battery life. The receiver system is also contained within a single printed circuit board (as opposed to the multi-component set-ups utilised within previous studies) which substantially reduces the overall weight and the likely mean time between failure (MTBF).

Previous UAVRTS have focused on incorporating and modifying either direction (e.g. direction of arrival) or range-based techniques (e.g. received signal strength) as methods of locating tags. We explored an alternative localisation method. Using grid flight mission functionality available in both open-source and commercial autopilot systems, we derived estimated tag locations by a simple mean coordinates calculation (Fig 1). The assumption of equal coverage of the surveyed area (provided by the flight grid), and the notion that the grid exceeds the range of the tag (i.e. so estimated locations are not simply the centre of the grid) are central to accurate tag location by this method. By flexibly altering the transmission power of the tag trigger exciter depending on the size of the grid employed, we ensure signal loss at grid edge regardless of the situation.

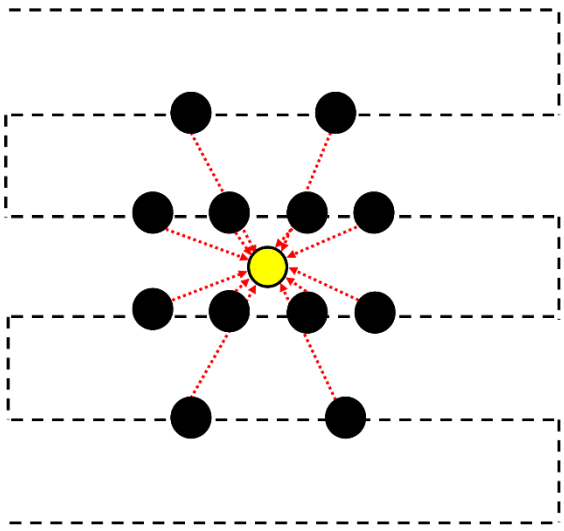


Figure 1. Locating tags using mean coordinates. Black line = UAV flight grid. Black dots = receiver location when tag transmission was detected. Yellow dot = estimated tag location based on mean coordinates of black dots.

To our knowledge, this method of localisation is undocumented for UAVRTS. We therefore sought to explore the level of accuracy deliverable, and the operational considerations that could affect it. We considered flight speed to be the key interest, exploring 1) the effect it had on the number of hits (tag responses) received and 2) how this affected the accuracy of a determined tag location. Beyond this, we then sought to test the real-world applicability of this system. Further objectives therefore included; 3) assessing the cross-compatibility of our UAVRTS to function on both multi-rotor and fixed-wing UAVs, and; 4) measuring the performance and reliability of the system with tags placed on animals.

Methods

RFID tag system design

The main components in each RFID tag (Fig.2) were a PIC10F206 microcontroller (Microchip Technology Inc, Chandler, Arizona, USA) and a HopeRF RFM69W radio transceiver (Hope Microelectronics co., Ltd, Nanshan District, Shenzhen, China) operating in the 868 MHz band. These were mounted on a custom-printed circuit board with integrated antenna and were powered by a single cell 60 mAh lithium polymer (LiPo) cell. The microcontroller was programmed to wake the radio module approximately once every 6 s for a period of approximately 2 ms. During this 2 ms period the radio module detected the signal from the trigger module (if one was present), switched to a different radio channel and responded using a simple medium access control delay mechanism with a radio packet containing a unique identifier for the tag. The response was transmitted three times, again using a simple medium access control delay mechanism to help reduce collisions between packets from different tags. In the absence of a signal, the microcontroller returned all components to a low power sleep state until the next listening window 6 s later. Predicted battery life in the absence of a trigger signal is over one year, but each transmission triggered will reduce the battery life by around one hour.. Responses from the tags were recorded by the UAV mounted receiver module which used a HopeRF RFM69W radio transceiver, a Quectel L86 GPS receiver (Quectel, Xuhui District, Shanghai, China) and an ATMega328P microcontroller (Microchip Technology Inc, Chandler, Arizona, USA). The microcontroller decoded the packets received from the RFID tag and saved the tag unique identifier, latitude and longitude of the receiver and time-stamp to a removable microSD memory card. With a clear line-of-sight and using a 10mW transmitter power in both directions (from trigger to tags and tags to receiver), the range achievable varied between 500 m and 800 m. Total weight for the system on the UAV was 195 g; this included; receiver box (115 g), trigger (80

g), batteries, and cable ties, etc. Each RFID tag weighed 9 g. At the time of writing, the estimated cost was £160 (£135 for receiver, £25 for trigger) with each RFID tag priced at just under £12.

Figure 2. Custom built radio-frequency identification tag.

UAV set up and RFID tag system integration

Multi-rotor: The platform was a DJI Phantom 3 professional (DJI, Nanshan District, Shenzhen, China). The receiver and tag trigger were mounted onto opposing ends of a 1 m long plastic rod, which in turn was cabled tied to the two landing stands on the drone (Fig. 3). The UAV was operated autonomously using the PIX4D capture app (PIX4D, Lausanne, Switzerland) on an Iphone 5S (Apple Inc., Cupertino, California, United States).

Fixed-wing: The UAV set-up was similar to that of (Ryan *et al.* 2015). The UAV airframe was a Skywalker X8 (Skywalker; Hubei, P.R.China). Autonomous flight capability was used (http://ardupilot.com/); this provided flight stabilisation, altitude control (including terrain

following), and GPS navigation. The tag trigger and receiver were located on opposite wing tips, with each accompanied by a single rechargeable 300 mAh LiPo cell as a power source, which could provide ~2 hours of use (Fig 3). The receiver was encased in a small plastic container wrapped with aluminium foil, except for directly above the GPS module, as initial testing revealed considerable radio interference from the fixed wing UAV avionics. Additional shielding was also fitted over the speed controller and electrical cables to the motor. The receiver case was bolted onto the wing tip, whilst the tag trigger was attached using cables ties, and the join further strengthened using cross-weave tape.

Figure 3. DJI Phantom 3 Pro with radio-frequency (RF) system mounted along a plastic rod attached to under carriage (top left). Skywalker X8 with RF system attached on wing tips (top right). Tag trigger mounted on X8 wing tip (bottom left). RF receiver mounted inside foil wrapped box on X8 wing tip (bottom right).

Static accuracy and the effect of UAV flight speed

Multi-rotor: Eighteen RFID tags were split into three groups which were each places at three different locations ~200 m apart (n = 6 per group), with every tag within each group equally spaced within a $1m^2$ area. Two GPS loggers (Ystumtec Ltd, Aberystwyth, United Kingdom) were present in each group to provide a reference location. The flight grid consisted of a four-line grid encompassing a 650 x 230 m area. Twelve flights were conducted in total at three different flight speeds based on percentage speed potentials of the DJI Phantom 3 pro, according to the PIX4D capture app, at 70% (~5.3 m/s), 80% (~8.5 m/s) and 100% (~14.5 m/s) of the maximum capable speed. Flight altitude was set to 100 m for all flights.

Fixed-wing: Fixed-wing UAVs are limited by their stall speed. In addition, wind speed affects performance, and thus it was impractical to attempt to test at varying speeds. Therefore, the accuracy of the tags was only assessed at a single target groundspeed set to 18 m/s. Twelve tags were placed (7 scattered; 5 within a 1 m^2 area) in a 3.69 ha field, and GPS location referenced (MyGPSCoordinates app; Kevin Willet, TappiApps) (reported accuracy +/-5m). A flight grid was created ~960 x 960 m (92 ha) in size, which included 19 lines at a spacing of 50 m, and flight altitude set to 100 m.

Attaching RFID tags to the sheep

Thirteen Herdwick sheep (*Ovis aries*) were selected for tag application. Ethical approval was obtained. The work described was conducted in accordance with the requirements of the UK Animals (Scientific Procedures) Act 1986, and with the approval of the Institute of Biological Environmental and Rural Sciences (IBERS) Animal Welfare and Ethical Review Board. Of the 13 sheep, two had a single tag attached to each of their horns, nine had a tag attached to one of their ears, and two had tags attached to collars fitted around their necks. In attaching

tags to the horns, tags were first dipped in IMPACT adhesive glue (Bostik Ltd, Stafford, United Kingdom), placed on top of the horn facing skywards, wrapped with a crepe bandage, then secured with a layer of RHINO cross weave fabric tape on top (Ultratape House, Dundee, UK). Each ear tag was secured to the outside edge of an existing 'loop' management tag using two cable ties. When attached to collars, tags were cable tied to the back of the collar facing upwards. The sheep were held in the same 3.69 ha field where the fixed UAV accuracy testing was undertaken. A similar flight grid was created (~960 x 960 m) and flight altitude set to 100 m. A total of 7 missions were completed over a two-week period.

Data analysis

Duplicate hits (as a result of the tag sending three responses per transmission) that shared the same position were deleted. Any duplicate responses that occurred after GPS update were treated as standalone responses as they had differing locations to the first response in the package. Only hits received along the grid lines were used, removing any that were recorded during launch/landing. Mean coordinates (Lat/Long) of each individual tag were subsequently calculated in open source GIS software (QGIS vers 2.12.3 *Lyon*). For assessing static accuracy, distance (in m) between each calculated tag mean coordinate (Lat2, Long2) and known GPS location (Lat1, Long1) was completed in Microsoft Excel using the following formula, which is based on the Spherical Law of Cosines:

$$acos\big(sin(lat1) * sin(lat2) + cos(lat1) * cos(lat2) * cos(long2 - long1)\big) * 6371$$

Statistical analyses were performed using R Studio version 3.6.1 (Team 2013). The packages MASS (Bates *et al.*, 2014) and lme4 (Venables & Ripley, 2002) were required. The 95th

percentile of the data was used as a measure of overall static accuracy.

Three regression analyses were performed. Firstly, in order to assess the relationship between accuracy (a positive, skewed, continuous variable), speed (continuous variable with values roughly close to 5 m/s, 8 m/s, 14 m/s and 18 m/s) and number of hits, gamma regression was used with speed and number of hits as explanatory variables. The functional form (i.e. whether higher order terms for hits was required) was motivated by local polynomial regression. Secondly, a (gamma) mixed effects model was used to provide an estimate of between-tag variability in accuracy. Thirdly, in order to assess the relationship between hits (an overdispersed count variable [mean = 21.84, variance = 400.82]) and speed, negative binomial regression was used. The Akaike information criterion (AIC) (comparing the model for hits predicted by speed, and a model for hits including only an intercept) was used to assess whether speed explained variability in the number of hits.

Results and Discussion

The R95 parameter was calculated to be 58.5 m (n =175, M=29.6 m, SE=1.46) (Fig. 4).

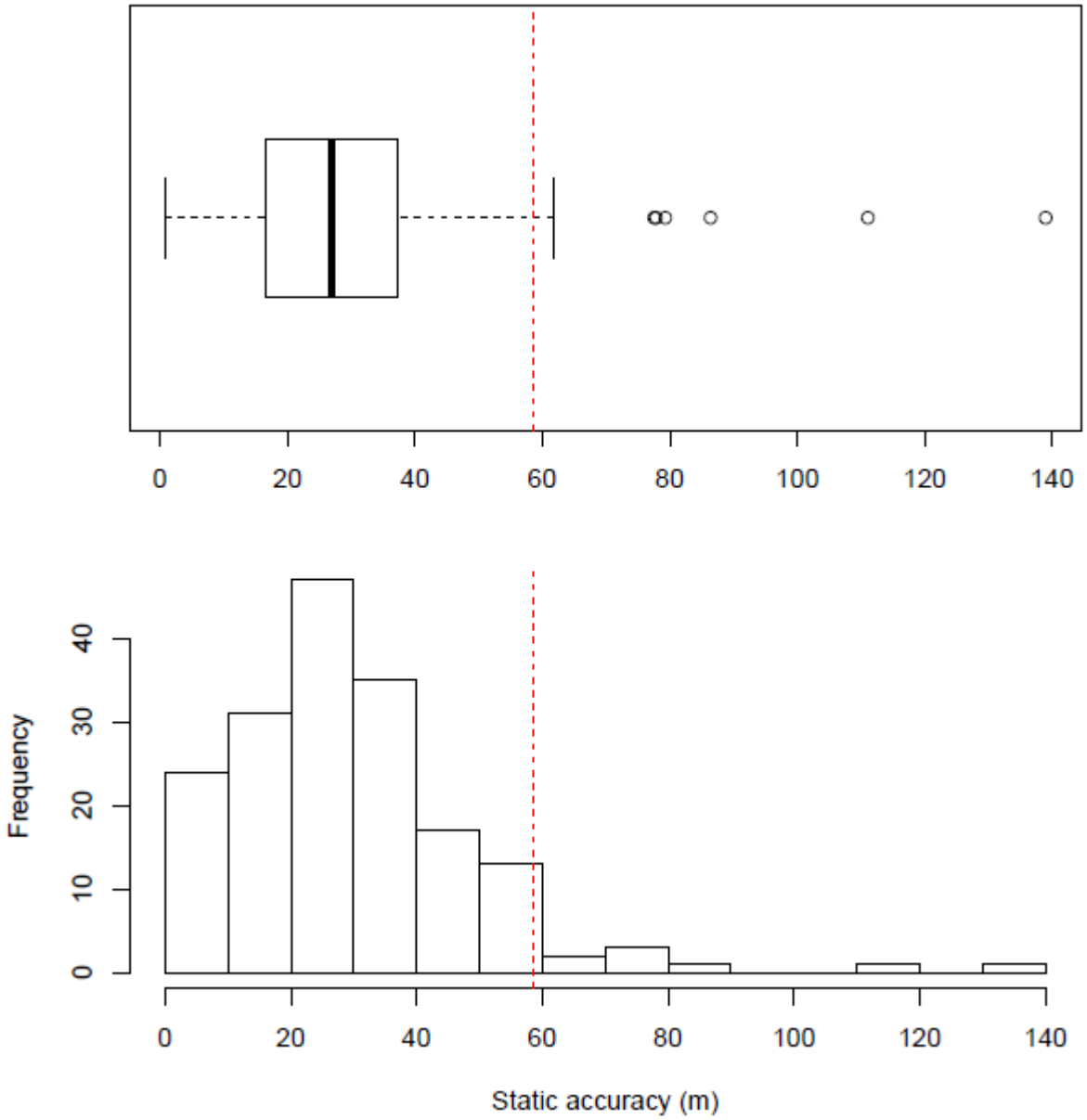


Figure 4. A histogram and boxplot of the static accuracy measurements, defined as the distance between each calculated tag mean coordinate produced from the RF system data, and known position of the RFID tags. The vertical red dotted line shows the R95 parameter with a value of 58.5 m.

Table 1. Gamma regression output. Variable hits was scaled and centred to avoid co-linearity issues. Only terms for hits were statistically significant.

	Estimate	SE	*t* value	*p* value
(Intercept)	3.11	0.12	25.81	<0.001
Speed	0.01	0.01	0.91	0.36
Hits	-8.9×10^{-3}	3.9×10^{-3}	-2.29	0.02
Hits Squared	3.7×10^{-4}	8.1×10^{-5}	4.63	<0.001

The multivariate analysis showed hits to be a more important variable in determining accuracy than speed: terms for hits were statistically significant (Table 1), whereas speed was not statistically significant after accounting for effects of variation in hits. However, speed will influence the number of hits (Table 2): higher speeds tend to result in fewer hits, as shown in Fig 5. Fig. 6 shows the observed relationship (as estimated by local polynomial regression) between static accuracy and number of hits, where lower values of accuracy imply better accuracy. Mean accuracy improves as the number of hits increases, but only up to ~25 hits. Thereafter, the mean accuracy declines for a larger number of hits.

Table 2. AIC scores for the negative binomial models for the number of hits. Including variable speed vastly improves model fit, suggesting speed explains variability in the number of hits.

Model	df	AIC
Intercept and Speed	3	1349.57
Only Intercept	2	1433.50

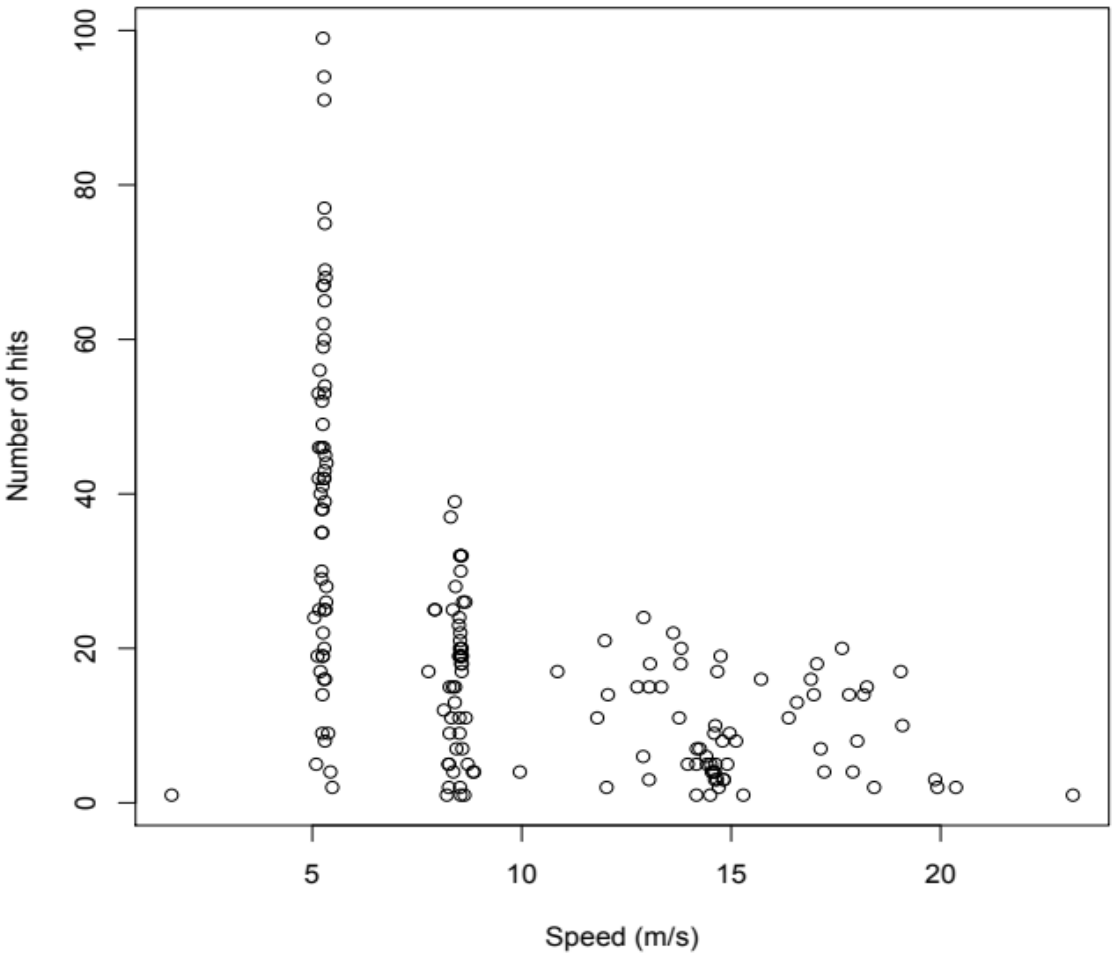


Figure 5. Effect of UAV speed (m/s) on observed number of hits (i.e. each successful package received from an RFID tag) by the RF tracking system.

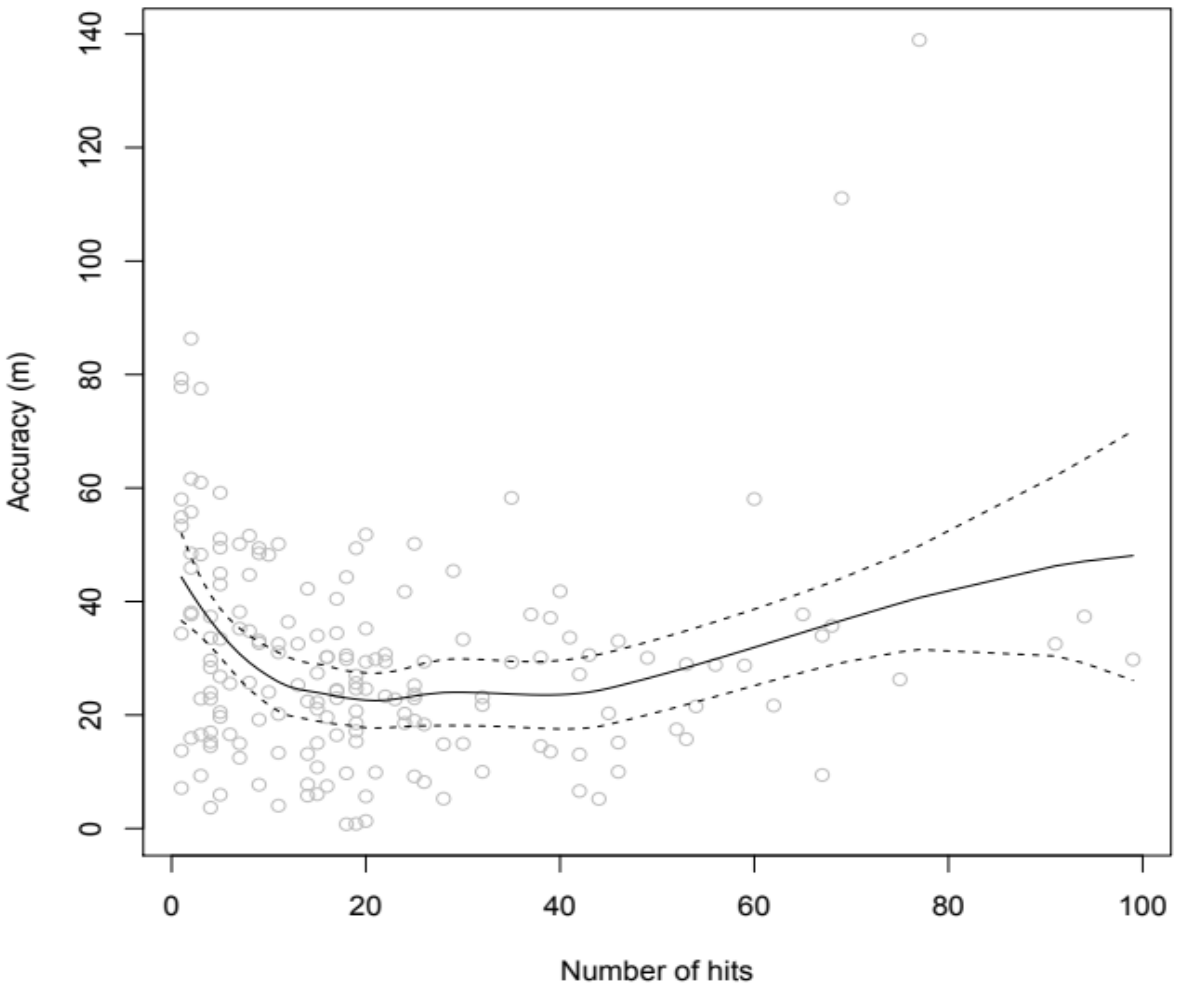


Figure 6. Effect of number of hits on observed static accuracy (i.e. distance between each calculated tag mean coordinate produced from the RF system data, and known position of the RFID tags) of the 175 data points (m). The black line shows the observed relationship between the mean accuracy and number of hits as estimated by local polynomial regression, and the dotted lines are 95% confidence intervals.

We reason that the variance witnessed in the data is likely due to component variation and suspected temperature compensation issues with the RF components, which remain to be fully quantified in future work. That said, including a random effect to assess variability between tags, the estimated between-ID tag variance was 0, indicating no between tag

variability (i.e. no individual tag was inherently more accurate than another). Overall, the comparatively high level accuracy achieved as a low-cost RF based system likely outweighs the observed variation in precision. Though further work is required, the results indicate that in considering static accuracy, increasing the number of hits (either by decreasing the time between pings transmitted on the tag trigger/or decreasing UAV speed) are key factors worth exploring.

All the RFID tags attached to the management ear tags of the sheep (n = 9) and one attached to the collars were still in place at the end of the experiment (n = 2), however none of those attached to the horns (n = 2) remained. The tags that fell off occurred before the first flight had been undertaken. Of the 10 still attached to the sheep, 9 worked with 100% consistency across all 7 recorded missions, while one failed after two missions. Tag response reliability was therefore calculated as 93%.

Since positioning is calculated from first to last response of each individual tag within the flight, only a time period and not a precise time point can be ascribed to calculated position, with the length of the time period depending on the size of the flight grid being performed. Used in conjunction with GPS systems (e.g. tracking collars), clarity could be improved. For example, GPS tracking of a small number of animals would provide a high number of consistent recordings over a time period, whereas the UAV based RF system could deliver a lower temporal number of 'snapshot' recordings for the entire group of animals.

When preparing flight grids, specifics such as line width are comparatively minor considerations relative to the need for grid to be large enough, and the tag trigger power to be low enough for the UAV to be able to fly out of range of the tags. A criticism of this system

could be that a rough radius of all combined tags in an area must be known in order to construct a grid to cover them all. When used on ungulates (who typically herd), or in conjunction with GPS loggers already on the ground, this may be more easily definable. Furthermore, the use of quick reconnaissance flights could be used to identify the spread of the target group. Another limitation is that although fixed-wing UAVs are capable of a large range the maximum grid size may be limited by the UAV operating regulations of the respective country. Though beyond visual line of sight (BVLOS) authorisation has the potential to extend the operational capability, this is still a developing framework in many countries and therefore may not be immediately accessible. Under current circumstances flying adjacent grids of a legal size in succession is a workable alternative.

Though we have demonstrated the viability of this system across multiple UAV platforms and provided considerations for its use, several key issues would benefit from further research and development. The system's performance in situations where animals may be in shaded/covered locations (e.g. woodlands, rocky areas) needs to be investigated before it is deployed in such circumstances. In addition, integration of the RF system with the UAV autopilot modules would allow more sophisticated surveying methods, such as circling or slowing down when a tag is detected in order to increase accuracy further.

Conclusions

This paper presents the first-cross platform compatible UAVRTS. Its flexibility and low-cost nature, together with the degree of accuracy achievable and proven ability to be utilised on mammals demonstrate its readiness as a field ready tool. Though applicable in many environments/situations, we contend that currently the suitable applications of this system would be; 1) the tracking of large ungulate herds, or 2) target animals which are located in enclosures, or have defined home range areas.

Acknowledgments

We would like to thank staff at Pwllpeiran Upland Research Centre for their help in the field, and I. Johnstone (Royal Society for the Protection of Birds (RSPB)) and I. Joyce (Elan Valley Trust) for observations and comments during development. Funding for the work was provided by Aberystwyth University's IBERS and Department of Computer Science, the RSPB and the Elan Valley Trust.

Author's contributions

BR, MN, and NS conceived the ideas and designed the methodology. BR, NS and FL collected the data; FL, TC and BR analysed the data. BR and MF led the writing of the manuscript. All authors contributed critically to the drafts and gave final approval for publication.

Data and system design accessibility

Data are R code used for the statistical analysis are available on the online open access repository Figshare (Roberts *et al.* 2020). Also included are logic diagrams of the ear tag hardware and software, and associated pseudocode. For further enquiries regarding system design please contact; ber32@aber.ac.uk or info@ystumtec.co.uk. The DOI is: https://doi.org/10.6084/m9.figshare.12712124.v3

References

Hodgson, J.C., Mott, R., Baylis, S.M., Pham, T.T., Wotherspoon, S., Kilpatrick, A.D., Raja Segaran, R., Reid, I., Terauds, A. & Koh, L.P. (2018) Drones count wildlife more accurately and precisely than humans. *Methods in Ecology and Evolution*, **9,** 1160-1167.

Kays, R., Crofoot, M.C., Jetz, W. & Wikelski, M. (2015) Terrestrial animal tracking as an eye on life and planet. *Science*, **348,** aaa2478.

Körner, F., Speck, R., Göktogan, A.H. & Sukkarieh, S. (2010) Autonomous airborne wildlife tracking using radio signal strength. *Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference on*, pp. 107-112. IEEE.

Ryan, J.C., Hubbard, A.L., Box, J.E., Todd, J., Christoffersen, P., Carr, J.R., Holt, T.O. & Snooke, N.A. (2015) UAV photogrammetry and structure from motion to assess calving dynamics at Store Glacier, a large outlet draining the Greenland ice sheet.

Tremblay, J.A., Desrochers, A., Aubry, Y., Pace, P. & Bird, D.M. (2017) A low-cost technique for radio-tracking wildlife using a small standard unmanned aerial vehicle. *Journal of Unmanned Vehicle Systems*, **5,** 102-108.